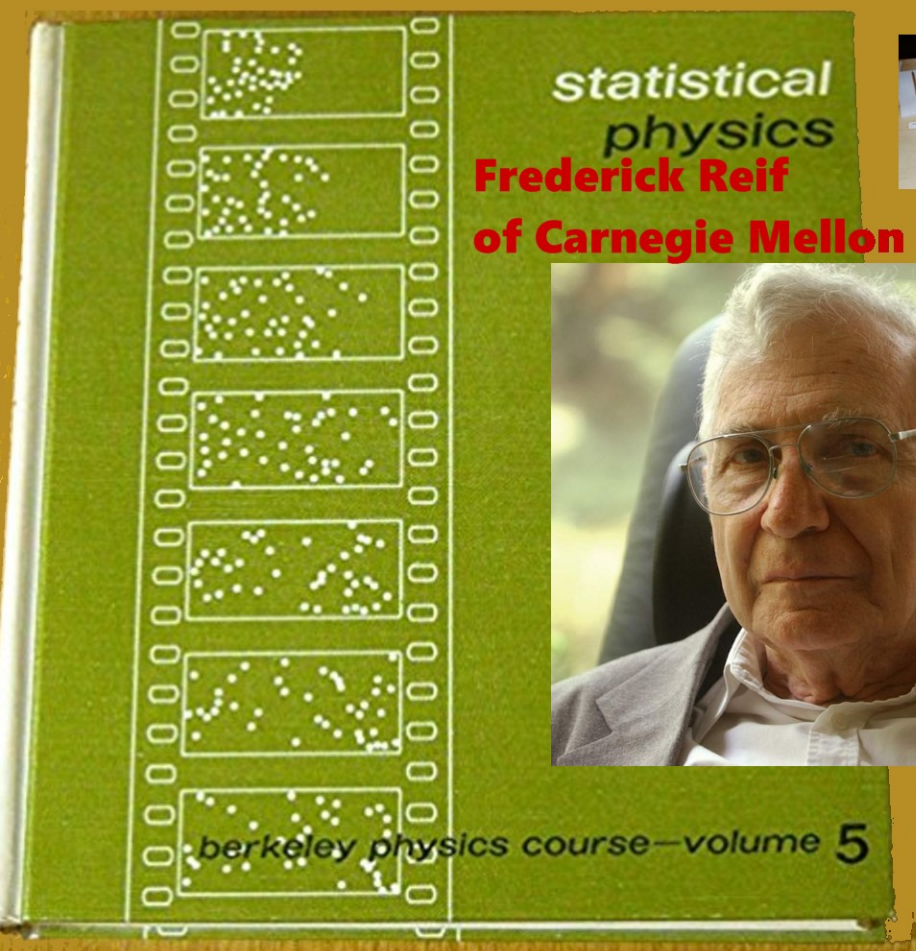
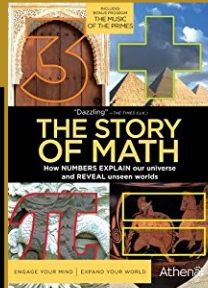
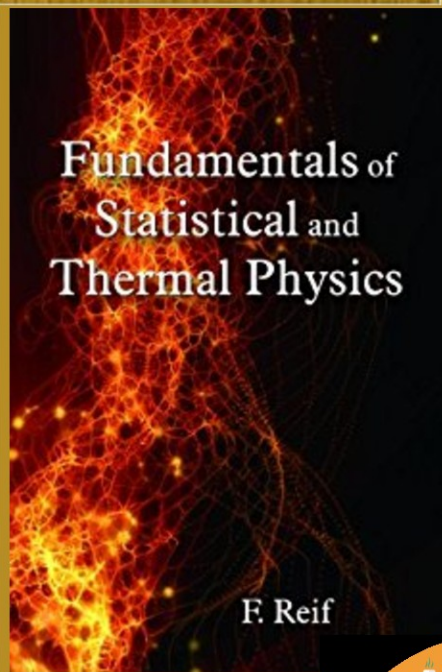
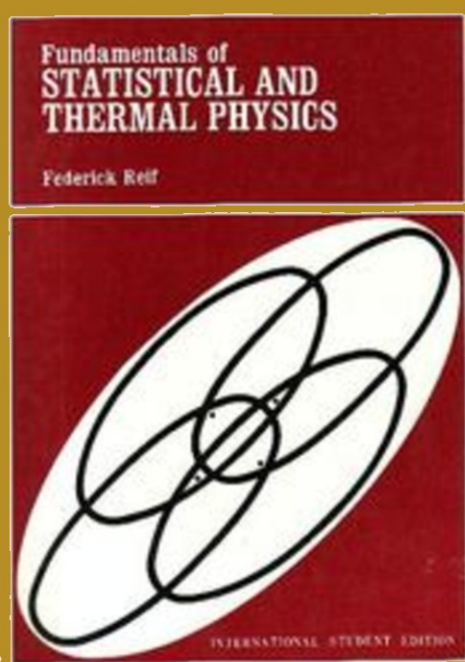
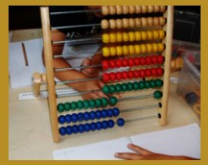


Batch PDF Merge

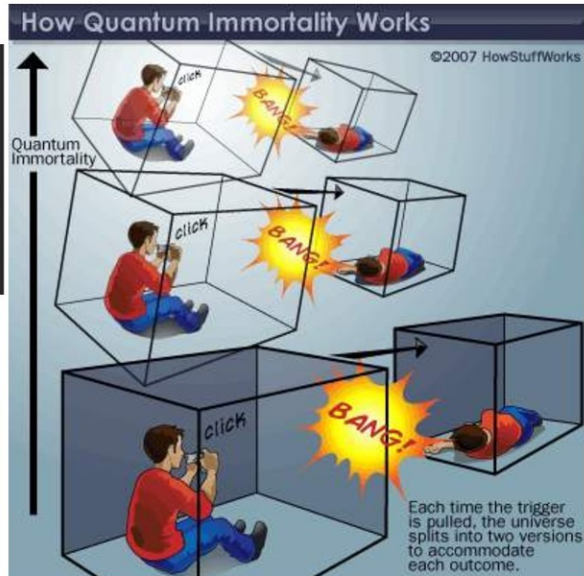
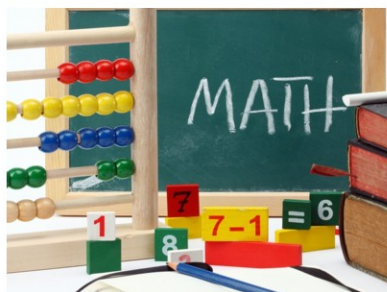
Batch PDF Merge



Frederick Reif
of Carnegie Mellon University



Fundamentals of Statistical and Thermal Physics



BBC
FOUR

PRECISION:

The Measure of All Things



vi *Preface to the Berkeley Physics Course*

The five volumes of the course as planned will include:

- I. Mechanics (Kittel, Knight, Ruderman)
- II. Electricity and Magnetism (Purcell)
- III. Waves (Crawford)
- IV. Quantum Physics (Wichmann)
- V. Statistical Physics (Reif)

The authors of each volume have been free to choose that style and method of presentation which seemed to them appropriate to their subject.

The initial course activity led Alan M. Portis to devise a new elementary physics laboratory, now known as the Berkeley Physics Laboratory. Because the course emphasizes the principles of physics, some teachers may feel that it does not deal sufficiently with experimental physics. The laboratory is rich in important experiments, and is designed to balance the course.

The financial support of the course development was provided by the National Science Foundation, with considerable indirect support by the University of California. The funds were administered by Educational Services Incorporated, a nonprofit organization established to administer curriculum improvement programs. We are particularly indebted to Gilbert Oakley, James Aldrich, and William Jones, all of ESI, for their sympathetic and vigorous support. ESI established in Berkeley an office under the very competent direction of Mrs. Mary R. Maloney to assist the development of the course and the laboratory. The University of California has no official connection with our program, but it has aided us in important ways. For this help we thank in particular two successive Chairmen of the Department of Physics, August C. Helmholz and Burton J. Moyer; the faculty and nonacademic staff of the Department; Donald Coney, and many others in the University. Abraham Olshen gave much help with the early organizational problems.

Your corrections and suggestions will always be welcome.

Eugene D. Commins
Frank S. Crawford, Jr.
Walter D. Knight
Philip Morrison
Alan M. Portis
Edward M. Purcell
Frederick Reif
Malvin A. Ruderman
Eyvind H. Wichmann
Charles Kittel, *Chairman*

January, 1965
Berkeley, California

Preface to Volume V

This last volume of the Berkeley Physics Course is devoted to the study of large-scale (i.e., *macroscopic*) systems consisting of many atoms or molecules; thus it provides an introduction to the subjects of statistical mechanics, kinetic theory, thermodynamics, and heat. The approach which I have followed is not patterned upon the historical development of these subjects and does not proceed along conventional lines. My aim has been + rather to adopt a modern point of view and to show, in as systematic and simple a way as possible, how the basic notions of atomic theory lead to a coherent conceptual framework capable of describing and predicting the properties of macroscopic systems.

In writing this book I have tried to keep in mind a student who, unencumbered by any prior familiarity with the subject matter, is encountering it for the first time from the vantage point of his previous study of elementary physics and atomic properties. I have therefore chosen an order of presentation which might seem well-motivated to such a student if he attempted to discover by himself how to gain an understanding of macroscopic systems. In seeking to make the presentation coherent and unified, I have based the entire discussion upon the systematic elaboration of a single principle, the tendency of an isolated system to approach its situation of greatest randomness. Although I have restricted my attention to simple systems, I have treated them by methods which are widely applicable and easily generalized. Above all, I have tried throughout to stress physical insight, the ability to perceive significant relationships quickly and simply. Thus I have attempted to discuss physical ideas at length without getting lost in mathematical formalism, to provide simple examples to illustrate abstract general concepts, to make numerical estimates of significant quantities, and to relate the theory to the real world of observation and experiment.

1 The subject matter to be covered in this volume had to be selected with great care. My intention has been to emphasize the most fundamental concepts which would be useful to physicists as well as to students of chemistry, biology, or engineering. The Teaching and Study Notes summarize the organization and contents of the book and provide some guides for the prospective teacher and student. The unconventional order of presentation, designed to stress the relation between the macroscopic and atomic levels of description, does not necessarily sacrifice the virtues inherent in

more traditional approaches. In particular, the following features may be worth mentioning:

(i) The student who completes Chap. 7 (even if he omits Chap. 6) will know the fundamental principles and basic applications of classical thermodynamics as well as if he had studied the subject along traditional lines. Of course, he will also have acquired much more insight into the meaning of entropy and a considerable knowledge of statistical physics.

(ii) I have been careful to emphasize that the statistical theory leads to certain results which are purely macroscopic in content and completely independent of whatever models one might assume about the atomic structure of the systems under consideration. The generality and model-independence of the classical thermodynamic laws is thus made explicitly apparent.

(iii) Although a historical approach rarely provides the most logical or illuminating introduction to a subject, an acquaintance with the evolution of scientific ideas is both interesting and instructive. I have, therefore, included in the text some pertinent remarks, references, and photographs of prominent scientists, all designed to give the student some perspective concerning the historical development of the subject.

The prerequisites needed for a study of this volume include, besides a rudimentary knowledge of classical mechanics and electromagnetism, only an acquaintance with the simplest atomic concepts and with the following quantum ideas in their most unsophisticated form: the meaning of quantum states and energy levels, Heisenberg's uncertainty principle, the de Broglie wavelength, the notion of spin, and the problem of a free particle in a box. The mathematical tools required do not go beyond simple derivatives and integrals, plus an acquaintance with Taylor's series. Any student familiar with the essential topics covered in the preceding volumes of the Berkeley Physics Course (particularly in Vol. IV) would, of course, be amply prepared for the present book. The book could be used equally well, however, for the last part of any other modern introductory physics course or for any comparable course at, or above, the level of second-year college students.

As I indicated at the beginning of this preface, my aim has been to penetrate the essence of a sophisticated subject sufficiently to make it seem simple, coherent, and easily accessible to beginning students. Although the goal is worth pursuing, it is difficult to attain. Indeed, the writing of this book was for me an arduous and lonely task that consumed an incredible amount of time and left me feeling exhausted. It would be some slight compensation to know that I had achieved my aim sufficiently well so that the book would be found useful.

Acknowledgments

I am grateful to Professor Allan N. Kaufman for reading the final manuscript critically and for always being willing to let me have the benefit of his opinions. Professors Charles Kittel and Edward M. Purcell made valuable comments concerning an earlier version of the first couple of chapters. Among graduate students, I should like to mention Richard Hess, who made many helpful observations about the preliminary edition of this volume, and Leonard Schlessinger, who worked out complete solutions for the problems and who provided the answers listed at the end of the book. I feel especially indebted to Jay Dratler, an undergraduate student who read both the preliminary edition and a substantial portion of the final manuscript. Starting out unfamiliar with the subject matter, he learned it himself from the book and exhibited in the process a fine talent for detecting obscurities and making constructive suggestions. He is probably the person who has contributed most significantly toward the improvement of the book.

The making of the computer-constructed pictures took an appreciable amount of time and effort. I wish, therefore, to express my warmest thanks to Dr. Berni J. Alder who helped me enormously in this task by personal cooperation uncontaminated by financial compensation. My ideas about these pictures could never have come to fruition if he had not put his computing experience at my disposal. We hope to continue our collaboration in the future by making available some computer-constructed movies which should help to illustrate the same ideas in more vivid form.

Mrs. Beverly Sykes, and later on Mrs. Patricia Cannady, were my loyal secretaries during the long period that I was occupied with this book. I feel very much indebted to them for their skill in deciphering and typing its successive handwritten versions. I also owe thanks to several other persons for their assistance in the production of the book. Among these are Mrs. Mary R. Maloney and Mrs. Lila Lowell, who were always willing to help with miscellaneous chores, and Mr. Felix Cooper, who is responsible for the execution of the artwork. Finally, I am grateful to Mr. William R. Jones, of Educational Services, Inc., for his efforts in handling relations with the National Science Foundation.

This volume owes an enormous debt to *Fundamentals of Statistical and Thermal Physics* (FSTP), my earlier book published by McGraw-Hill in 1965, which represented an attempt at educational innovation at the more advanced level of upper-division college students. My extensive experience derived from FSTP, and many details of presentation, have been incorporated in the present volume.† I wish, therefore, to express my gratitude to

† Some portions of the present volume are thus also subject to the copyright provisions of *Fundamentals of Statistical and Thermal Physics*.

those individuals who assisted me during the writing of FSTP as well as to those who have provided me with constructive criticisms since its publication. I should also like to thank the McGraw-Hill Book Company for disregarding the conflicting copyright provisions to give me unrestricted permission to include material from FSTP in the present volume. Although I am not dissatisfied with the general approach developed in FSTP, I have come to recognize that the exposition there could often have been simpler and more penetrating. I have, therefore, made use of these new insights to include in the present volume all the improvements in organization and wording intended for a second edition of FSTP. By virtue of its similar point of view, FSTP may well be a useful reference for students interested in pursuing topics beyond the level of the present book; such students should, however, be cautioned to watch out for certain changes of notation.

Although the present volume is part of the Berkeley Physics Course project, it should be emphasized that the responsibility for writing this volume has been mine alone. If the book has any flaws (and I myself am aware of some even while reading proof), the onus for them must, therefore, rest upon my own shoulders.

Teaching and Study Notes

Organization of the Book

The book is divided into three main parts which I shall describe in turn:

Part A: Preliminary Notions (Chapters 1 and 2)

Chapter 1: This chapter provides a qualitative introduction to the most fundamental physical concepts to be explored in this book. It is designed to make the student aware of the characteristic features of macroscopic systems and to orient his thinking along fruitful lines.

Chapter 2: This chapter is somewhat more mathematical in nature and is intended to familiarize the student with the basic notions of probability theory. No prior knowledge of probability ideas is assumed. The concept of ensembles is stressed throughout and all examples are designed to illuminate physically significant situations. Although this chapter is oriented toward subsequent applications in the remainder of the book, the probability concepts discussed are, of course, intended to be useful to the student in far wider contexts.

These chapters should not take too much time. Indeed, some students may well have sufficient background preparation to be familiar with some of the material in these chapters. Nevertheless, I would definitely recommend that such students *not* skip these two chapters, but that they consider them a useful review.

Part B: Basic Theory (Chapters 3, 4, and 5)

This part constitutes the heart of the book. The logical and quantitative development of the subject of this volume really starts with Chapter 3. (In this sense the first two chapters could have been omitted, but this would have been very unwise pedagogically.)

Chapter 3: This chapter discusses how a system consisting of many particles is described in statistical terms. It also introduces the basic postulates of the statistical theory. By the end of this chapter the student should have come to realize that the quantitative understanding of macroscopic systems hinges essentially on considerations involving the counting of the states accessible to the systems. He may not yet perceive, however, that this insight has much useful value.

Chapter 4: This chapter constitutes the real pay-off. It starts out, innocently enough, by investigating how two systems interact by heat transfer alone. This investigation leads very quickly, however, to the fundamental concepts of entropy, of absolute temperature, and of the canonical distribution (or Boltzmann factor). By the end of the chapter the student is in a position to deal with thoroughly practical problems—indeed, he has learned how to calculate from first principles the paramagnetic properties of a substance or the pressure of an ideal gas.

Chapter 5: This chapter brings the ideas of the theory completely down to earth. Thus it discusses how one relates atomic concepts to macroscopic measurements and how one determines experimentally such quantities as absolute temperature or entropy.

An instructor thoroughly pressed for time can stop at the end of these five chapters without too many pangs of conscience. At this point a student should have acquired a fairly good understanding of absolute temperature, entropy, and the Boltzmann factor—i.e., of the most fundamental concepts of statistical mechanics and thermodynamics. (Indeed, the only thermodynamic result still missing concerns the fact that the entropy remains unchanged in a quasi-static adiabatic process.) At this stage I would consider the minimum aims of the course to have been fulfilled.

Part C: Elaboration of the Theory (Chapters 6, 7, and 8)

This part consists of three chapters which are independent of each other in the sense that any one of them can be taken up without requiring the others as prerequisites. In addition, it is perfectly possible to cover only the first few sections of any one of these chapters before turning to another of these chapters. Any instructor can thus use this flexibility to suit his own predilections or the interests of his students. Of the three chapters, Ch. 7 is the one of greatest fundamental importance in rounding out the theory; since it completes the discussion of thermodynamic principles, it is also the one likely to be most useful to students of chemistry or biology.

Chapter 6: This chapter discusses some particularly important applications of the canonical distribution by introducing approximate classical notions into the statistical description. The Maxwell velocity distribution of molecules in a gas and the equipartition theorem are the main topics of this chapter. Illustrative applications include molecular beams, isotope separation, and the specific heat of solids.

Chapter 7: This chapter begins by showing that the entropy remains unchanged in a process which is adiabatic and quasi-static. This completes the discussion of the thermodynamic laws which are then summarized in

their full generality. The chapter then proceeds to examine a few important applications: general equilibrium conditions, including properties of the Gibbs free energy; equilibrium between phases; and implications for heat engines and biological organisms.

Chapter 8: This last chapter is intended to illustrate the discussion of the nonequilibrium properties of a system. It treats the transport properties of a dilute gas by the simplest mean-free-path arguments and deals with viscosity, thermal conductivity, self-diffusion, and electrical conductivity.

This completes the description of the essential organization of the book. In the course as taught at Berkeley, the aim is to cover the major part of this book in about eight weeks of the last quarter of the introductory physics sequence.

The preceding outline should make clear that, although the presentation of the subject matter of the book is unconventional, it is characterized by a tight logical structure of its own. This logical development may well seem more natural and straightforward to the student, who approaches the topics without any preconceptions, than to the instructor whose mind is molded by conventional ways of teaching the subject. I would advise the instructor to think the subject through afresh himself. If sheer force of habit should lead him to inject traditional points of view injudiciously, he may disrupt the logical development of the book and thus confuse, rather than enlighten, the student.

Other Features of the Book

Appendix: The four sections of the appendix contain some peripheral material. In particular, the Gaussian and Poisson distributions are specifically discussed because they are important in so many diverse fields and because they are also relevant in the laboratory part of the Berkeley Physics Course.

Mathematical Notes: These notes constitute merely a collection of mathematical tidbits found useful somewhere in the text or in some of the problems.

Mathematical Symbols and Numerical Constants: These can be found listed at the end of the book and also on its inside front and back covers.

Summaries of Definitions: These are given at the ends of chapters for ease of reference and convenience of review.

Problems: The problems constitute a very important part of the book. I have included about 160 of them to provide an ample and thought-provoking collection to choose from. Although I would not expect a student to

work through all of them, I would encourage him to solve an appreciable fraction of the problems at the end of any chapter he has studied; otherwise he is likely to derive little benefit from the book. Problems marked by stars are somewhat more difficult. The supplementary problems deal mostly with material discussed in the appendices.

Answers to Problems: Answers to most of the problems are listed at the end of the book. The availability of these answers should facilitate the use of the book for self-study. Furthermore, although I would recommend that a student try to solve each problem before looking at the answer given for it, I believe that it is pedagogically valuable if he can check his own answer immediately after he has worked out a solution. In this way he can become aware of his mistakes early and thus may be stimulated to do further thinking instead of being lulled into unjustified complacency. (Although I have tried to assure that the answers listed in the book are correct, I cannot guarantee it. I should appreciate being informed of any mistakes that might be uncovered.)

Subsidiary material: Material which consists of illustrations or various remarks is set in two-column format with smaller type in order to differentiate it from the main skeleton of logical development. Such subsidiary material should not be skipped in a first reading, but might be passed over in subsequent reviews.

Equation numbering: Equations are numbered consecutively within each chapter. A simple number, such as (8), refers to equation number 8 in the chapter under consideration. A double number is used to refer to equations in other chapters. Thus (3.8) refers to Equation (8) in Chap. 3, (A.8) to Equation (8) in the Appendix, (M.8) to Equation (8) in the Mathematical Notes.

+ *Advice to the Student*

Learning is an active process. Simply reading or memorizing accomplishes practically nothing. Treat the subject matter of the book as though you were trying to discover it yourself, using the text merely as a guide that you should leave behind. The task of science is to learn ways of thinking which are effective in describing and predicting the behavior of the observed world. The only method of learning new ways of thinking is to practice thinking. Try to strive for insight, to find new relationships and simplicity where before you saw none. Above all, do not simply memorize formulas; learn modes of reasoning. The *only* relations worth remembering deliberately are the few Important Relations listed explicitly at the end of

- each chapter. If these are not sufficient to allow you to reconstruct in your head any other significant formula in about twenty seconds or less, you have not understood the subject matter.
- † Finally, it is much more important to master a few fundamental concepts than to acquire a vast store of miscellaneous facts and formulas. If in the text I have seemed to belabor excessively some simple examples, such as the system of spins or the ideal gas, this has been deliberate. It is particularly true in the study of statistical physics and thermodynamics that some apparently innocent statements are found to lead to remarkable conclusions of unexpected generality. Conversely, it is also found that many problems can easily lead one into conceptual paradoxes or seemingly hopeless calculational tasks; here again, a consideration of simple examples can often resolve the conceptual difficulties and suggest new calculational procedures or approximations. Hence my last advice is that you try to understand simple basic ideas well and that you then proceed to work many problems, both those given in the book and those resulting from questions you may pose yourself. Only in this way will you test your understanding and learn how to become an independent thinker in your own right.

Contents

Preface to the Berkeley Physics Course v

Preface to Volume V vii

Acknowledgments ix

Teaching and Study Notes xi

54 Chapter 1 *Characteristic Features of Macroscopic Systems* 1

- 1.1 Fluctuations in Equilibrium 4
- 1.2 Irreversibility and the Approach to Equilibrium 15
- 1.3 Further Illustrations 29
- 1.4 Properties of the Equilibrium Situation 31
- 1.5 Heat and Temperature 35
- 1.6 Typical Magnitudes 39
- 1.7 Important Problems of Macroscopic Physics 45
 - Summary of Definitions 50
 - Suggestions for Supplementary Reading 51
 - Problems 51

44 Chapter 2 *Basic Probability Concepts* 55

- 2.1 Statistical Ensembles 56
- 2.2 Elementary Relations among Probabilities 64
- 2.3 The Binomial Distribution 67
- 2.4 Mean Values 75
- 2.5 Calculation of Mean Values for a Spin System 80
- 2.6 Continuous Probability Distributions 86
 - Summary of Definitions 90
 - Important Relations 90
 - Suggestions for Supplementary Reading 91
 - Problems 91

42 Chapter 3 Statistical Description of Systems of Particles 99

- 3.1 Specification of the State of a System 101
- 3.2 Statistical Ensemble 108
- 3.3 Statistical Postulates 111
- 3.4 Probability Calculations 116
- + 3.5 Number of States Accessible to a Macroscopic System 118
- 3.6 Constraints, Equilibrium, and Irreversibility 124
- 3.7 Interaction between Systems 129
 - Summary of Definitions 135
 - Important Relations 136
 - Suggestions for Supplementary Reading 136
 - Problems 136

50 Chapter 4 Thermal Interaction 141

- 4.1 Distribution of Energy between Macroscopic Systems 142
- 4.2 The Approach to Thermal Equilibrium 147
- 4.3 Temperature 149
- 4.4 Small Heat Transfer 155
- 4.5 System in Contact with a Heat Reservoir 157
- 4.6 Paramagnetism 163
- 4.7 Mean Energy of an Ideal Gas 166
- 4.8 Mean Pressure of an Ideal Gas 172
 - Summary of Definitions 176
 - Important Relations 177
 - Suggestions for Supplementary Reading 177
 - Problems 178

32 *Chapter 5 Microscopic Theory and Macroscopic Measurements 191*

- 5.1 Determination of the Absolute Temperature 192
- 5.2 High and Low Absolute Temperatures 196
- 5.3 Work, Internal Energy, and Heat 200
- 5.4 Heat Capacity 206
- 5.5 Entropy 209
- 5.6 Intensive and Extensive Parameters 211
 - Summary of Definitions 213
 - Important Relations 213
 - Suggestions for Supplementary Reading 213
 - Problems 214

42 *Chapter 6 Canonical Distribution in the Classical Approximation 223*

- 6.1 The Classical Approximation 224
- 6.2 Maxwell Velocity Distribution 231
- 6.3 Discussion of the Maxwell Distribution 235
- 6.4 Effusion and Molecular Beams 240
- 6.5 The Equipartition Theorem 246
- 6.6 Applications of the Equipartition Theorem 248
- 6.7 The Specific Heat of Solids 250
 - Summary of Definitions 256
 - Important Relations 256
 - Suggestions for Supplementary Reading 256
 - Problems 257

52 Chapter 7 *General Thermodynamic Interaction* 265

- 7.1 Dependence of the Number of States on the External Parameters 266
- 7.2 General Relations Valid in Equilibrium 271
- 7.3 Applications to an Ideal Gas 276
- 7.4 Basic Statements of Statistical Thermodynamics 281
- 7.5 Equilibrium Conditions 286
- 7.6 Equilibrium between Phases 292
- 7.7 The Transformation of Randomness into Order 299
 - Summary of Definitions 307
 - Important Relations 307
 - Suggestions for Supplementary Reading 308
 - Problems 308

32 Chapter 8 *Elementary Kinetic Theory of Transport Processes* 317

- 8.1 Mean Free Path 319
- 8.2 Viscosity and Transport of Momentum 323
- 8.3 Thermal Conductivity and Transport of Energy 331
- 8.4 Self-diffusion and Transport of Molecules 335
- 8.5 Electrical Conductivity and Transport of Charge 339
 - Summary of Definitions 342
 - Important Relations 342
 - Suggestions for Supplementary Reading 343
 - Problems 343

Appendix 349

- A.1 Gaussian Distribution 350
- A.2 Poisson Distribution 355
- A.3 Magnitude of Energy Fluctuations 357
- A.4 Molecular Impacts and Pressure in a Gas 360

Mathematical Notes 363

- M.1 The Summation Notation 364
- M.2 Sum of a Geometric Series 364
- M.3 Derivative of $\ln n!$ for large n 365
- M.4 Value of $\ln n!$ for large n 366
- M.5 The Inequality $\ln x \leq x - 1$ 367
- M.6 Evaluation of the Integral $\int_{-\infty}^{\infty} e^{-x^2} dx$ 367
- M.7 Evaluation of Integrals of the Form $\int_{-\infty}^{\infty} e^{-ax^2} x^n dx$ 369

*Supplementary Problems 371**Mathematical Symbols 377**Greek Alphabet 379**Numerical Constants 381**Answers to Problems 383**Index 393*

Chapter 1

Characteristic Features of Macroscopic Systems

1.1	Fluctuations in Equilibrium	4
1.2	Irreversibility and the Approach to Equilibrium	15
1.3	Further Illustrations	29
1.4	Properties of the Equilibrium Situation	31
1.5	Heat and Temperature	35
1.6	Typical Magnitudes	39
1.7	Important Problems of Macroscopic Physics	45
	Summary of Definitions	50
	Suggestions for Supplementary Reading	51
	Problems	51

Chapter 1 Characteristic Features of Macroscopic Systems

Dass ich erkenne, was die Welt
Im Innersten zusammenhält,
Schau' alle Wirkenskraft und Samen,
Und tu' nicht mehr in Worten kramen.

Goethe, *Faust*†

The entire world of which we are aware through our senses consists of objects that are *macroscopic*, i.e., large compared to atomic dimensions and thus consisting of very many atoms or molecules. This world is enormously varied and complex, encompassing gases, liquids, solids, and biological organisms of the most diverse forms and compositions. Accordingly, its study has formed the subject matter of physics, chemistry, biology, and several other disciplines. In this book we want to undertake the challenging task of gaining some insights into the fundamental properties of all macroscopic systems. In particular, we should like to investigate how the few unifying concepts of atomic theory can lead to an understanding of the observed behavior of macroscopic systems, how quantities describing the directly measurable properties of such systems are interrelated, and how these quantities can be deduced from a knowledge of atomic characteristics.

Scientific progress made in the first half of this century has led to very basic knowledge about the structure of matter on the *microscopic* level, i.e., on the small scale of the order of atomic size (10^{-8} cm). Atomic theory has been developed in quantitative detail and has been buttressed by an overwhelming amount of experimental evidence. Thus we know that all matter consists of molecules built up of atoms which, in turn, consist of nuclei and electrons. We also know the quantum laws of microscopic physics governing the behavior of atomic particles. Hence we should be in a good position to exploit these insights in discussing the properties of macroscopic objects.

Indeed, let us justify this hope in greater detail. Any macroscopic system consists of very many atoms. The laws of quantum mechanics describing the dynamical behavior of atomic particles are well established. The electromagnetic forces responsible for the interactions between these atomic particles are also very well understood.

† From Faust's opening soliloquy in Goethe's play, Part I, Act I, Scene I, lines 382-385. Translation: "That I may recognize what holds the world together in its inmost essence, behold the driving force and source of everything, and rummage no more in empty words."

Ordinarily they are the only forces relevant because gravitational forces between atomic particles are usually negligibly small compared to electromagnetic forces. In addition, a knowledge of nuclear forces is usually not necessary since the atomic nuclei do not get disrupted in most ordinary macroscopic physical systems and in all chemical and biological systems.† Hence we can conclude that our knowledge of the laws of microscopic physics should be quite adequate to allow us, in principle, to deduce the properties of any macroscopic system from a knowledge of its microscopic constituents.

It would, however, be quite misleading to stop on this optimistic note. A typical macroscopic system of the type encountered in everyday life contains about 10^{25} interacting atoms. Our concrete scientific aim is that of understanding and predicting the properties of such a system on the basis of a minimum number of fundamental concepts. We may well know that the laws of quantum mechanics and electromagnetism describe completely all the atoms in the system, whether it be a solid, a liquid, or a human being. But this knowledge is utterly useless in achieving our scientific aim of prediction unless we have available methods for coping with the tremendous complexity inherent in such systems. The difficulties involved are not of a type which can be solved merely by the brute force application of bigger and better electronic computers. The problem of 10^{25} interacting particles dwarfs the capabilities of even the most fanciful of future computers; furthermore, unless one asks the right questions, reams of computer output tape are likely to provide no insight whatever into the essential features of a problem. It is also worth emphasizing that complexity involves much more than questions of quantitative detail. In many cases it can lead to remarkable qualitative features that may seem quite unexpected. For instance, consider a gas of identical simple atoms (e.g., helium atoms) which interact with each other through simple known forces. It is by no means evident from this microscopic information that such a gas can condense very abruptly so as to form a liquid. Yet this is precisely what happens. An even more striking illustration is provided by any biological organism. Starting solely from a knowledge of atomic structure, would one suspect that a few simple kinds of atoms, forming certain types of molecules, can give rise to systems capable of biological growth and self-reproduction?

The understanding of macroscopic systems consisting of very many particles thus requires primarily the formulation of new concepts capable of dealing with complexity. These concepts, based ultimately

† Gravitational and nuclear interactions may, however, become pertinent in some astrophysical problems.

- difficult to see, notice,
or understand, esp.
with difficulty.
- (readily) with no
difficulty.

(4)-

+ ▸

(4)

(5)

(i)

(ii)

upon the known fundamental laws of microscopic physics, should achieve the following aims: make apparent the parameters most useful in describing macroscopic systems; permit us to discern readily the essential characteristics and regularities exhibited by such systems; and finally, provide us with relatively simple methods capable of predicting quantitatively the properties of these systems.

The discovery of concepts sufficiently powerful to achieve these aims represents clearly a major intellectual challenge, even when the fundamental laws of microscopic physics are assumed known. It is thus not surprising that the study of complex systems consisting of many atoms occupies much attention in research at the forefront of physics. On the other hand, it is remarkable that quite simple reasoning is sufficient to lead to substantial progress in the understanding of macroscopic systems. As we shall see, the basic reason is that the very presence of a large number of particles allows one to use statistical methods with singular effectiveness.

It is by no means obvious how we should go about achieving our aim of understanding macroscopic systems. Indeed, their apparent complexity may seem forbidding. In setting out on our path of discovery we shall, therefore, follow good scientific procedure by first examining some simple examples. At this stage we shall not let our imagination be stifled by trying to be rigorous or overly critical. Our purpose in this chapter is rather to recognize the essential features characteristic of macroscopic systems, to see the main problems in qualitative outline, and to get some feeling for typical magnitudes. This preliminary investigation should serve to suggest to us appropriate methods for attacking the problems of macroscopic systems in a systematic and quantitative way.

1.1 Fluctuations in Equilibrium

A simple example of a system consisting of many particles is a gas of identical molecules, e.g., argon (Ar) or nitrogen (N_2) molecules. If the gas is dilute (i.e., if the number of molecules per unit volume is small), the average separation between the molecules is large and their mutual interaction is correspondingly small. The gas is said to be ideal if it is sufficiently dilute so that the interaction between its molecules is almost negligible.† An ideal gas is thus particularly simple. Each of its molecules spends most of its time moving like a free particle unin-

† The interaction is "almost" negligible if the total potential energy of interaction between the molecules is negligible compared to their total kinetic energy, but is sufficiently large so that the molecules can interact and thus exchange energy with each other.

fluenced by the presence of the other molecules or the container walls; only rarely does it come sufficiently close to the other molecules or the container walls so as to interact (or *collide*) with them. In addition, if the gas is sufficiently dilute, the average separation between its molecules is much larger than the average de Broglie wavelength of a molecule. In this case quantum-mechanical effects are of negligible importance and it is permissible to treat the molecules as distinguishable particles moving along classical trajectories.†

Consider then an ideal gas of N molecules confined within a container or box. In order to discuss the simplest possible situation, suppose that this whole system is *isolated* (i.e., that it does not interact with any other system) and that it has been left undisturbed for a very long time. We now imagine that we can observe the gas molecules, without affecting their motion, by using a suitable camera to take a motion picture of the gas. Successive frames of the film would show the positions of the molecules at regular intervals separated by some short time τ_0 . We then could examine the frames individually, or alternatively, could run the movie through a projector.

In the latter case we would observe on the projection screen a picture showing the gas molecules in constant motion: Thus any given molecule moves along a straight line until it collides with some other molecule or with the walls of the box; it then continues moving along some other straight line until it collides again; and so on and so forth. Each molecule moves strictly in accordance with the laws of motion of mechanics. Nevertheless, N molecules moving throughout the box and colliding with each other represent a situation so complex that the picture on the screen appears rather *chaotic* (unless N is very small).

Let us now focus attention on the positions of the molecules and their distribution in space. To be precise, consider the box to be divided by some imaginary partition into two equal parts (see Fig. 1.1). Denote the number of molecules in the left half of the box by n and the number in the right half by n' . Of course

$$n + n' = N, \quad (1)$$

the total number of molecules in the box. If N is large, we would ordinarily find that $n \approx n'$, i.e., that roughly half of the molecules are in each half of the box. We emphasize, however, that this statement is only approximately true. Thus, as the molecules move throughout the box, colliding occasionally with each other or with the walls, some

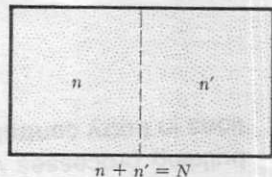


Fig. 1.1 A box containing an ideal gas consisting of N molecules. The box is shown subdivided into two equal parts by an imaginary partition. The number of molecules in the left half is denoted by n , the number of molecules in the right half by n' .

† The validity of the classical approximation will be examined more extensively in Sec. 6.3.

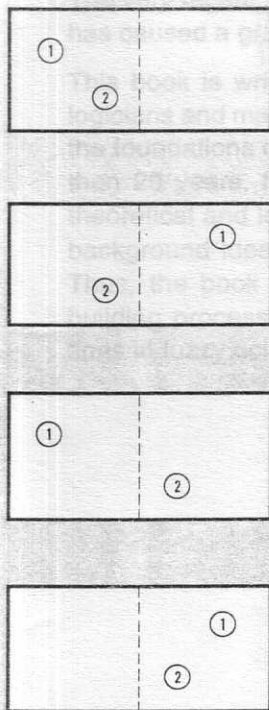


Fig. 1.2 Schematic diagram illustrating the 4 different ways in which 2 molecules can be distributed between the two halves of a box.

of them enter the left half of the box, while others leave it. Hence the number n of molecules actually located in the left half fluctuates constantly in time (see Figs. 1.3 through 1.6). Ordinarily these fluctuations are small enough so that n does not differ too much from $\frac{1}{2}N$. There is, however, nothing that prevents all molecules from being in the left half of the box (so that $n = N$, while $n' = 0$). Indeed this *might* happen. But how likely is it that it actually *does* happen?

To gain some insight into this question, let us ask in how many ways the molecules can be distributed between the two halves of the box. We shall call each distinct way in which the molecules can be distributed between these two halves a *configuration*. A single molecule can then be found in the box in two possible configurations, i.e., it can be either in the left half or in the right half. Since the two halves have equal volumes and are otherwise equivalent, the molecule is equally likely to be found in either half of the box.† If we consider 2 molecules, each one of them can be found in either of the 2 halves. Hence the total number of possible configurations (i.e., the total number of possible ways in which the 2 molecules can be distributed between the two halves) is equal to $2 \times 2 = 2^2 = 4$ since there are, for each possible configuration of the first molecule, 2 possible configurations of the other (see Fig. 1.2). If we consider 3 molecules, the total number of their possible configurations is equal to $2 \times 2 \times 2 = 2^3 = 8$ since there are, for each of the 2^2 possible configurations of the first 2 molecules, 2 possible configurations of the last one. Similarly, if we consider the general case of N molecules, the total number of possible configurations is $2 \times 2 \times \dots \times 2 = 2^N$. These configurations are listed explicitly in Table 1.1 for the special case where $N = 4$.

Note that there is only one way of distributing the N molecules so that all N of them are in the left half of the box. It represents only one special configuration of the molecules compared to the 2^N possible configurations of these molecules. Hence we would expect that, among a very large number of frames of the film, on the average only one out of every 2^N frames would show all the molecules to be in the left half. If P_N denotes the fraction of frames showing all the N molecules located in the left half of the box, i.e., if P_N denotes the relative frequency, or *probability*, of finding all the N molecules in the left half, then

$$P_N = \frac{1}{2^N}. \quad (2)$$

† We assume that the likelihood of finding a particular molecule in any half of the box is unaffected by the presence there of any number of other molecules. This will be true if the total volume occupied by the molecules themselves is negligibly small compared to the volume of the box.

1	2	3	4	n	n'	C(n)
L	L	L	L	4	0	1
L	L	L	R	3	1	4
L	L	R	L	3	1	
L	R	L	L	3	1	
R	L	L	L	3	1	
L	L			2	2	6
L		L		2	2	
L			L	2	2	
	L	L		2	2	
	L		L	2	2	
		L	L	2	2	
L	R	R	R	1	3	4
R	L	R	R	1	3	
R	R	L	R	1	3	
R	R	R	L	1	3	
R	R	R	R	0	4	1

Table 1.1 Enumeration of the 16 possible ways in which $N = 4$ molecules (denoted by 1, 2, 3, 4) can be distributed between two halves of a box. The letter *L* indicates that the particular molecule is in the left half of the box, the letter *R* that it is in the right half. The number of molecules in each of the halves is denoted by n and n' , respectively. The symbol $C(n)$ denotes the number of possible configurations of the molecules when n of them are in the left half of the box.

The probability density $= 1/2^N = p_n$
is related to 2^N

Notice (between two halves)
hence the 2 appearing
in 2^N , i.e. there is more
than 2 parts (m) $\rightarrow$
 m^N

Consider when m is
very large $\rightarrow \infty$ $\rightarrow$ 2^N
 $m(\Delta V) = x = 6$
 $\Delta V = 0.2$ of 0.2 the
volume of particle occupies

Volume $= V = m(\Delta V)$

(Complete)

Computer-constructed pictures

The following pages and several subsequent ones show figures constructed by means of a high-speed electronic digital computer. The situation investigated in every case is the classical motion of several particles in a box, the particles being represented by disks moving in two dimensions. The forces between any two particles, or between a particle and a wall, are assumed to be like those between "hard" objects (i.e., to vanish when they do not touch and to become infinite when they do touch). All resulting collisions are thus elastic. The computer is given some initial specified positions and velocities of the particles. It is then asked to solve numerically the equations of motion of these particles for all subsequent (or prior) times and

to display pictorially on a cathode-ray oscilloscope the positions of the molecules at successive times $t = j\tau_0$ where τ_0 is some small fixed time interval and where $j = 0, 1, 2, 3, \dots$. A movie camera photographing the oscilloscope screen then yields the successive picture frames reproduced in the figures. (The time interval τ_0 was chosen long enough so that several molecular collisions occur between the successive frames displayed in the figures.) The computer is thus used to simulate in detail a hypothetical experiment involving the dynamical interaction between many particles.

All the computer-made pictures were produced with the generous cooperation of Dr. B. J. Alder of the Lawrence Radiation Laboratory at Livermore.

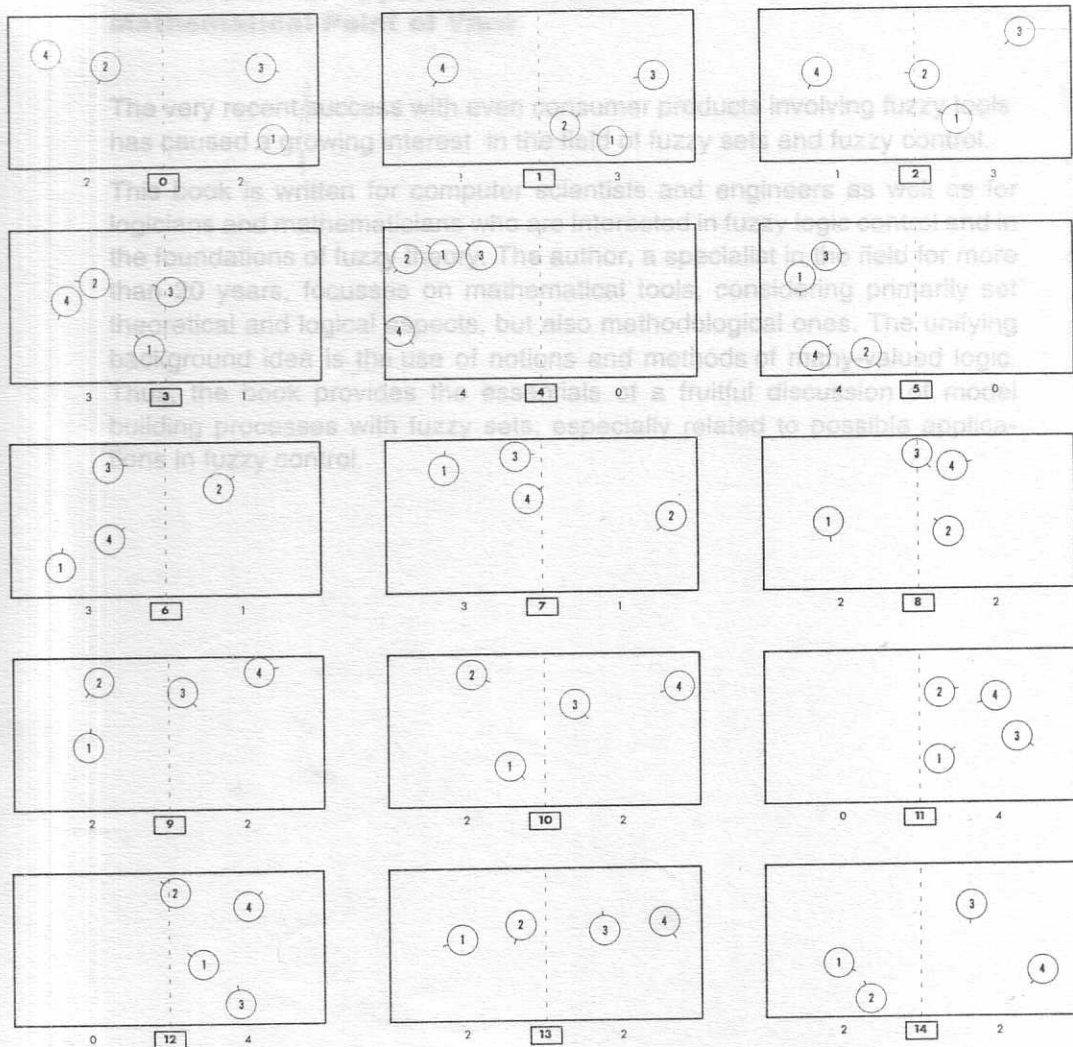


Fig. 1.3 Computer-made pictures showing 4 particles in a box. The fifteen successive frames (labeled by $j = 0, 1, 2, \dots, 14$) are pictures taken a long time after the beginning of the computation with assumed initial conditions. The number of particles

located in each half of the box is printed directly beneath that half. The short line segment emanating from each particle indicates the direction of the particle's velocity.

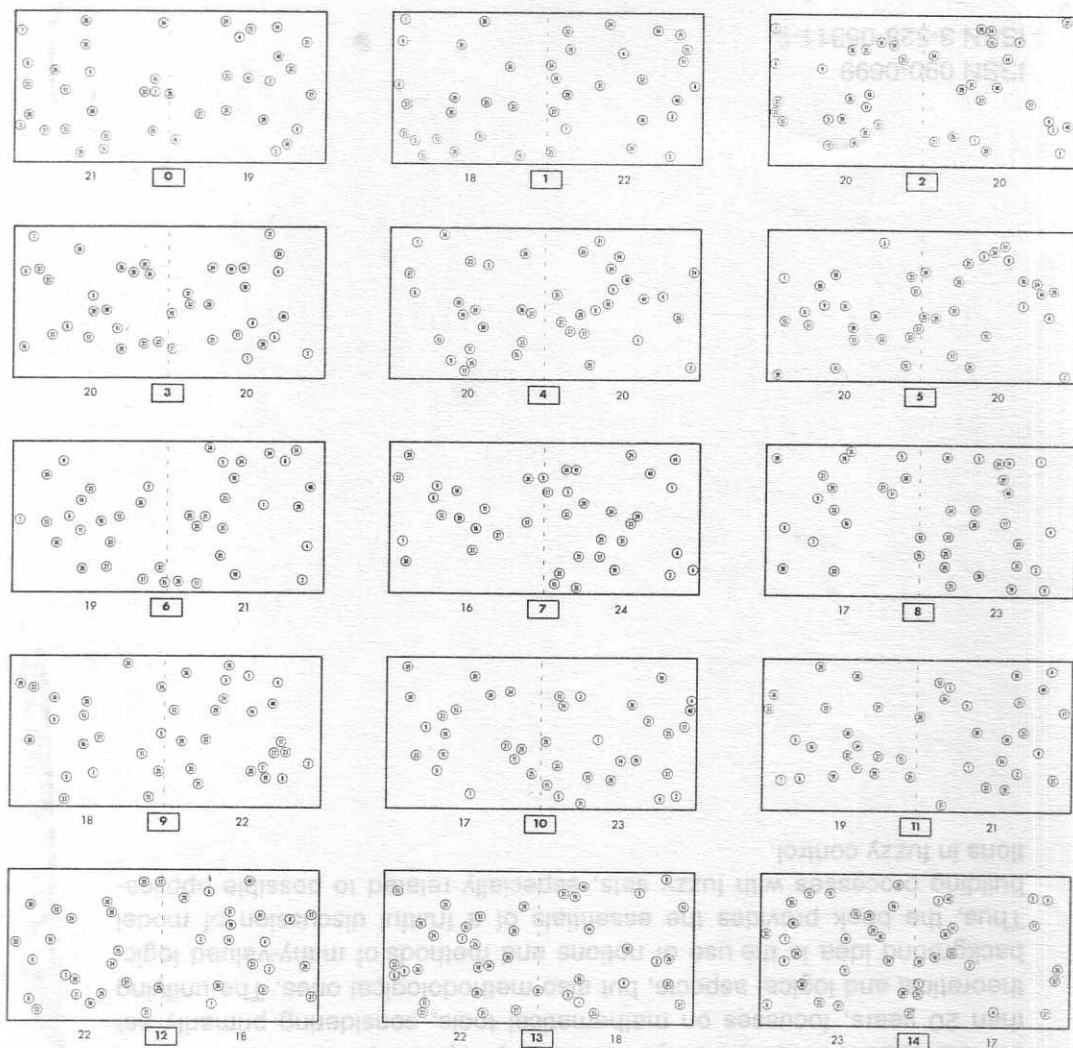


Fig. 1.4 Computer-made pictures showing 40 particles in a box. The fifteen successive frames (labeled by $j = 0, 1, 2, \dots, 14$) are pictures taken a long time after the beginning of the compu-

tation with assumed initial conditions. The number of particles located in each half of the box is printed directly beneath that half. The velocities of the particles are not indicated.

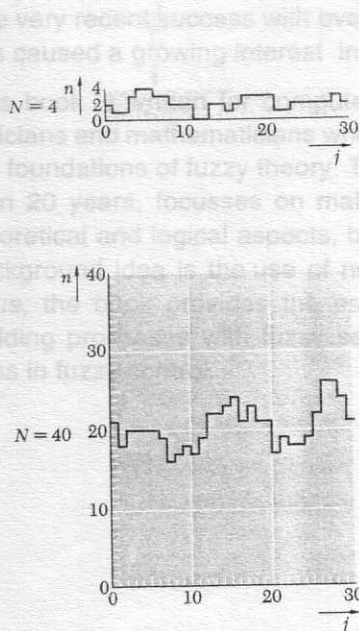


Fig. 1.5 The number n of particles in the left half of the box as a function of the frame index j or the elapsed time $t = j\tau_0$. The number n in the j th frame is indicated by a horizontal line extending from j to $j + 1$. The graphs describe Fig. 1.3 for $N = 4$ particles and Fig. 1.4 for $N = 40$ particles, but contain information about more frames than were shown there.

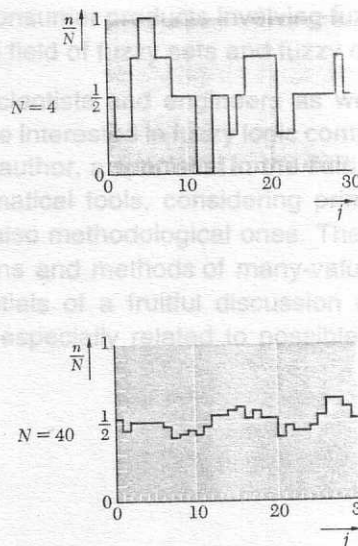


Fig. 1.6 The relative number n/N of particles in the left half of the box as a function of the frame index j or the elapsed time $t = j\tau_0$. The information presented is otherwise the same as that in Fig. 1.5.

Similarly, the case where no molecule at all is in the left half is also very special since there is again only one such configuration of the molecules out of 2^N possible configurations. Thus the probability P_0 of finding no molecule located in the left half should also be given by

$$P_0 = \frac{1}{2^N}. \quad (3)$$

More generally, consider a situation where n of the N molecules of the gas are located in the left half of the box and let us denote by $C(n)$ the number of possible configurations of the molecules in this case. [That is, $C(n)$ is the number of possible ways the molecules can be distributed in the box so that n of them are found in the left half of the box.] Since the total number of possible configurations of the molecules is 2^N , one would expect that, among a very large number of frames of the film, on the average $C(n)$ out of every 2^N such frames would show n molecules to be in the left half of the box. If P_n denotes the fraction of frames showing n molecules located in the left half, i.e., if P_n denotes the relative frequency, or probability, of finding n molecules in the left half, then

$$P_n = \frac{C(n)}{2^N}. \quad (4)$$

Example

Consider the special case where the gas consists of only four molecules. The number $C(n)$ of possible configurations of each kind is listed in Table 1.1. Suppose that a movie of this gas consists of a great many frames. Then we expect that the fraction P_n of these frames showing n molecules in the left half (and cor-

respondingly $n' = N - n$ molecules in the right half) is given by:

$$P_4 = P_0 = \frac{1}{16},$$

$$P_3 = P_1 = \frac{4}{16} = \frac{1}{4}, \quad (4a)$$

$$P_2 = \frac{6}{16} = \frac{3}{8}.$$

As we have seen, a situation where $n = N$ (or where $n = 0$) corresponds only to a single possible molecular configuration. More generally, if N is large, then $C(n) \ll 2^N$ if n is even moderately close to N (or even moderately close to 0). In other words, a situation where the distribution of molecules is so nonuniform that $n \gg \frac{1}{2}N$ (or that $n \ll \frac{1}{2}N$) corresponds to relatively few configurations. A situation of this kind, which can be obtained in relatively few ways, is rather special and is accordingly said to be relatively *nonrandom* or *orderly*;

• Indistinguishable w.r to $C(n)$
• but distinguishable w.r to
the number the particle
carries.

according to (4), it occurs relatively infrequently. On the other hand, a situation where the distribution of the molecules is almost uniform, so that $n \approx n'$, corresponds to many possible configurations; indeed, as is illustrated in Table 1.1, $C(n)$ is maximum if $n = n' = \frac{1}{2}N$. A situation of this kind, which can be obtained in many different ways, is said to be *random* or *disordered*; according to (4) it occurs quite frequently. In short, more random (or uniform) distributions of the molecules in the gas occur more frequently than less random ones. The physical reason is clear: All molecules must move in a very special way if they are to concentrate themselves predominantly in one part of the box; similarly, if they are all located in one part of the box, they must move in a very special way if they are to remain concentrated there.

The preceding statements can be made quantitative by using Eq. (4) to calculate the actual probability of occurrence of a situation where any number n of molecules are in the left half of the box. We shall postpone until the next chapter the requisite computation of the number of molecular configurations $C(n)$ in the general case. It is, however, very easy and illuminating to consider an extreme case and ask how frequently one would expect all the molecules to be located in the left half of the box. Indeed, (2) asserts that a fluctuation of this kind would, on the average, be observed to occur only once in every 2^N frames of the film.

To gain an appreciation for magnitudes, we consider some specific examples. If the gas consisted of only 4 molecules, all of them would, on the average, be found in the left half of the box once in every 16 frames of the film. A fluctuation of this kind would thus occur with moderate frequency. On the other hand, if the gas consisted of 80 molecules, all of these would, on the average, be found in the left half of the box in only one out of $2^{80} \approx 10^{24}$ frames of the film. This means that, even if we took a million pictures every second, we would have to run the film for a time appreciably greater than the age of the universe before we would have a reasonable chance of obtaining one frame showing all the molecules in the left half of the box.† Finally, suppose that we consider as a realistic example a box having a volume of 1 cm^3 and containing air at atmospheric pressure and room temperature. Such a box contains about 2.5×10^{19} molecules. [See Eq. (27) later in this chapter.] A fluctuation where all of these are located in one half of the box would, on the average, appear in only one out of

† There are about 3.15×10^7 seconds in a year and the estimated age of the universe is of the order of 10^{10} years.

$$22.5 \times 10^{19} \approx 107.5 \times 10^{18}$$

frames of the film. (This number of frames is so fantastically large that it could not be accumulated even though our film ran for times incredibly larger than the age of the universe.) Fluctuations where not all, but a predominant majority of the molecules are found in one half of the box, would occur somewhat more frequently; but this frequency of occurrence would still be exceedingly small. Hence we arrive at the following general conclusion: *If the total number of particles is large, fluctuations corresponding to an appreciably nonuniform distribution of these molecules occur almost never.*

Let us now conclude by summarizing our discussion of the isolated ideal gas which has been left undisturbed for a long time. The number n of molecules in one half of the box fluctuates in time about the constant value $\frac{1}{2}N$ which occurs most frequently. The frequency of occurrence of a particular value of n decreases rapidly the more n differs from $\frac{1}{2}N$, i.e., the greater the difference $|\Delta n|$ where

$$\Delta n \equiv n - \frac{1}{2}N. \quad (5)$$

Indeed, if N is large, only values of n with $|\Delta n| \ll N$ occur with significant frequency. Positive and negative values of Δn occur equally often. The time dependence of n has thus the appearance indicated schematically in Fig. 1.7.

The gas can be described in greatest detail by specifying its *microscopic state*, or *microstate*, at any time, i.e., by specifying the maximum possible information about the gas molecules at this time (e.g., the position and velocity of each molecule). From this microscopic point of view, a hypothetical film of the gas appears very complex since the locations of the individual molecules are different in every frame of the film. As the individual molecules move about, the microscopic state of the gas changes thus in a most complicated way. From a large-scale or *macroscopic point of view*, however, one is not interested in the behavior of each and every molecule, but in a much less detailed description of the gas. Thus the *macroscopic state*, or *macrostate*, of the gas at any time might be quite adequately described by specifying merely the *number* of molecules located in any part of the box at this time.† From this macroscopic point of view, the isolated gas which



Fig. 1.7 Schematic illustration showing how the number n of molecules in one half of a box fluctuates as a function of the time t . The total number of molecules is N .

† To be specific, we could imagine that the box is subdivided into many equal cells, each having a volume large enough to contain many molecules ordinarily. The macroscopic state of the gas could then be described by specifying the number of molecules located in each such cell.

has been left undisturbed for a long time represents a very simple situation since its macroscopic state does not tend to change in time. Indeed, suppose that, starting at some time t_1 , we observed the gas over some moderately long period of time τ by taking a movie of it. Alternatively, suppose that, starting at some other time t_2 , we again observed the gas over the same time τ by taking a movie of it. From a macroscopic point of view these two movies would ordinarily look indistinguishable. In each case the number n of particles in the left half of the box would ordinarily fluctuate about the same value $\frac{1}{2}N$, and the magnitude of the observed fluctuations would ordinarily also look alike. Disregarding very exceptional occurrences (to be discussed in the next section), the observed macrostate of the gas is thus independent of the starting time of our observations; i.e., we can say that the macrostate of our gas does not tend to change in time. In particular, the value about which n fluctuates (or, more precisely, its *average* value) does not tend to change in time. A system of many particles (such as our gas) whose macroscopic state does not tend to change in time is said to be in *equilibrium*.

Remark

- To define the concept of a time average in precise terms, let us denote by $n(t)$ the number of molecules in the left half of the box at any time t . The time average value of n at any time t , taken over the time interval τ , can then be denoted by $[\bar{n}(t)]_\tau$ and is defined by

$$[\bar{n}(t)]_\tau \equiv \frac{1}{\tau} \int_t^{t+\tau} n(t') dt'. \quad (6)$$

- Equivalently, if a movie strip beginning at the time t extends over a time τ and contains $g = \tau/\tau_0$ frames occurring at successive times $t_1 = t, \quad t_2 = t + \tau_0, \quad t_3 = t + 2\tau_0, \dots, t_g = t + (g-1)\tau_0$, the definition (6) becomes

$$[\bar{n}(t)]_\tau = \frac{1}{g} [n(t_1) + n(t_2) + \dots + n(t_g)].$$

If we omit explicit indication of the time interval τ to be considered, $\bar{n}(t)$ implies an average over some appropriately chosen time interval τ of appreciable length. In the equilibrium situation of our gas, $\bar{n}$ tends to be constant and equal to $\frac{1}{2}N$.

1.2 Irreversibility and the Approach to Equilibrium

Consider an isolated gas consisting of a large number N of molecules. If the fluctuations occurring in this gas in equilibrium are such that n is ordinarily very close to its most probable value $\frac{1}{2}N$, under what conditions can we expect to find situations where n differs very appreciably from $\frac{1}{2}N$? Such situations can occur in two different ways which we shall discuss in turn.

Rarely occurring large fluctuations in equilibrium

Although in the gas in equilibrium n is ordinarily always close to $\frac{1}{2}N$, values of n far from $\frac{1}{2}N$ can occur, but they do so quite rarely. If we observe the gas long enough, we may thus succeed in observing, at some particular instant of time t , a value of n appreciably different from $\frac{1}{2}N$.

Suppose that such a large spontaneous fluctuation of $|\Delta n|$ has occurred, e.g., that n at some particular time t_1 assumes a value n_1 which is much greater than $\frac{1}{2}N$. What can we then say about the probable behavior of n as time goes on? To the extent that $|n_1 - \frac{1}{2}N|$ is large, the particular value n_1 corresponds to a highly nonuniform distribution of the molecules and occurs very rarely in equilibrium. It is then most likely that the value n_1 occurs as a result of a fluctuation which is represented by a peak whose maximum is near n_1 (as indicated by the peak marked X in Fig. 1.8). The reason is the following: It might be possible for a value as large as n_1 to occur also as a result of a fluctuation represented by a peak whose maximum is larger than n_1 (such as the peak marked Y in Fig. 1.8); but the occurrence of such a large fluctuation is much less likely still than the already rare occurrence of a smaller fluctuation such as X . Thus we may conclude that it is indeed most likely that the time t_1 , where $n = n_1$, corresponds to a peak (such as X) where n is maximum. The general behavior of n as a function of time is then, however, apparent from Fig. 1.8. As time proceeds, n must tend to decrease (with accompanying small fluctuations) until it reverts to the usual equilibrium situation where it no longer tends to change, but fluctuates merely about the constant average value $\frac{1}{2}N$. The approximate time required for the large fluctuation (where $n = n_1$) to decay back to the equilibrium situation (where $n \approx \frac{1}{2}N$) is called the *relaxation time* for decay of this fluctuation. Note that, among many movie strips of duration τ , a movie strip showing the gas near the time t_1 where a large fluctuation occurs is highly exceptional. Not only does such a movie strip occur very rarely but, if it does occur, it

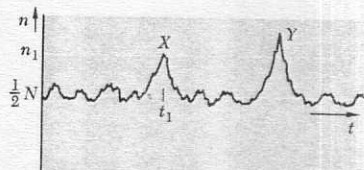


Fig. 1.8 Schematic illustration showing rare instances where the number n of molecules in one half of a box exhibits large fluctuations about its equilibrium value $\frac{1}{2}N$.

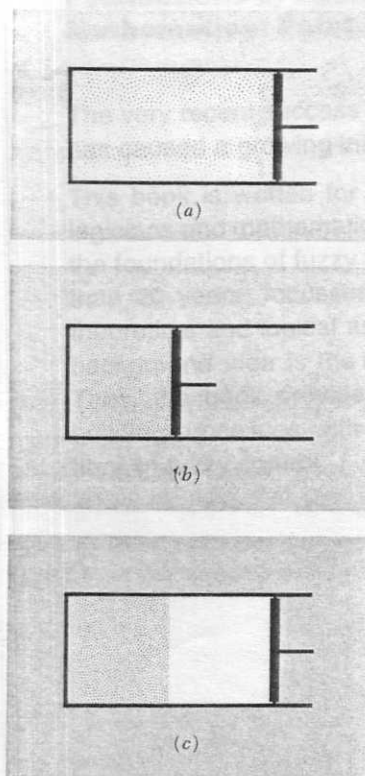


Fig. 1.9 The piston in (a) is moved to the position (b) so as to compress the gas into the left half of the box. When the piston is suddenly restored to its initial position, as shown in (c), the molecules immediately afterward are all located in the left half of the box, while the right half is empty.

is distinguishable from other movie strips since it shows a situation which tends to change in time.†

Our comments can thus be summarized in the following way: If n is known to assume a value n_1 which is appreciably different from its equilibrium average value $\frac{1}{2}N$, then n will almost† always change in such a direction as to approach the equilibrium value $\frac{1}{2}N$. In more physical terms, the value n_1 corresponds to a very nonuniform distribution of the molecules and the molecules would have to move in a very special way in order to preserve such nonuniformity. The incessant motion of the molecules thus almost always results in mixing them up so thoroughly that they become distributed over the entire box in the most random (or uniform) possible way. (See Figs. 1.15 through 1.20 at the end of this section.)

Remarks

Note that the statements of the preceding paragraph are equally applicable whether the large fluctuation ($n_1 - \frac{1}{2}N$) is positive or negative. If it is positive, the value n_1 will almost always correspond to the maximum of a fluctuation in n (such as that indicated by the peak X in Fig. 1.8). If it is negative, it will almost always correspond to the minimum of a fluctuation in n . The argument leading to the conclusion of the paragraph remains, however, essentially identical.

Note also that the statements of the paragraph remain equally valid irrespective of whether the change in time is in the forward or backward direction (i.e., irrespective of whether a movie of the gas is played forward or backward through a projector). If n_1 corresponds to a maximum such as that indicated by the peak X at time t_1 , then n must decrease both for $t > t_1$ and for $t < t_1$.

because the fluctuation is spontaneous

† This does not contradict the statement that the gas is in equilibrium in the long run, i.e., when observed over a very long time which may contain several occurrences of fluctuations as large as n_1 .

† We use the qualifying word "almost" since, instead of corresponding to a maximum of a peak such as X , the value n_1 may very rarely lie on the rising side of a peak such as Y . In this case n would initially increase, i.e., change away from its equilibrium value $\frac{1}{2}N$.

Initial Conditions (I.C.)

(2) Specially prepared initial situations

Although a nonrandom situation, where n is appreciably different from $\frac{1}{2}N$, may occur as a result of a spontaneous fluctuation of the gas in equilibrium, such a large fluctuation occurs so rarely that it would almost never be observed in practice. [Remember the numerical estimates based on Eq. (2) or (3).] Most macroscopic systems with which we deal, however, have not remained isolated and undisturbed for very long periods of time and, accordingly, are not in equilibrium. Non-random situations thus occur quite commonly, not as a result of spontaneous fluctuations of a system in equilibrium, but as a result of interactions which affected the system at some not too distant time in the past. Indeed, it is quite easy to bring about a nonrandom situation of a system by means of external intervention.

Examples

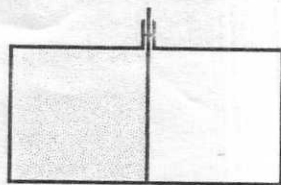
When a wall of a box is made movable, it becomes a *piston*. One can use such a piston (as shown in Fig. 1.9) to compress a gas into the left half of a box. When the piston is suddenly restored to its initial position, all the molecules immediately afterward are still in the left half of the box. Thus one has produced an extremely nonuniform distribution of the molecules in the box.

Equivalently, consider a box divided into two equal parts by a partition (see Fig. 1.10). Its left half is filled with N molecules of a gas, while its right half is empty. If the gas is in equilibrium under these circumstances, the distribution of its molecules is essentially uniform throughout the left half of the box. Imagine that the partition is now suddenly removed. The molecules immediately afterward are then still uniformly distributed throughout the left half of the box. This distribution, however, is highly non-uniform under the new conditions which leave the molecules free to move throughout the entire box.

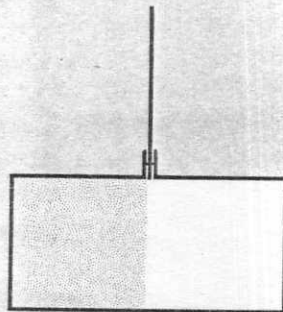
w.r.t. I.C. ($t \rightarrow 0$)

boundary conditions
has changed.

not in equilibrium
w.r.t. the whole box
ie new I.C. ($t \rightarrow \infty$)



(a)



(b)

Fig. 1.10 When the partition in (a) is suddenly removed, all the molecules immediately afterward are located in the left half of the box, as shown in (b).

Suppose that an isolated system is known to be in a highly nonrandom situation, e.g., suppose that all the molecules in a gas are known to be predominantly in the left half of the box so that n is appreciably different from $\frac{1}{2}N$. Then it is essentially irrelevant whether the system got into this condition by virtue of a very rare spontaneous fluctuation in equilibrium, or whether it got there by virtue of some form of prior external intervention. Irrespective of its past history, the subsequent behavior of the system in time will thus be similar to that discussed previously when we considered the decay of a large fluctuation in equilibrium. In short, since nearly all possible ways in which the molecules of the system can move will result in a more random distribution of these molecules, the situation of the system will almost always tend to change in time so as to become as random as possible. After the most random condition has been attained, it exhibits then no further tendency to change, i.e., it represents the ultimate equilibrium condition attained by the system. For example, Fig. 1.11 indicates schematically what happens after the partition of Fig. 1.10 is suddenly removed. The number n of molecules in the left half tends to change from its initial value $n = N$ (corresponding to a highly nonuniform distribution of the molecules in the box) until the ultimate equilibrium situation is attained where $n \approx \frac{1}{2}N$ (corresponding to an essentially uniform distribution of molecules in the box). (See Figs. 1.12 and 1.18.)

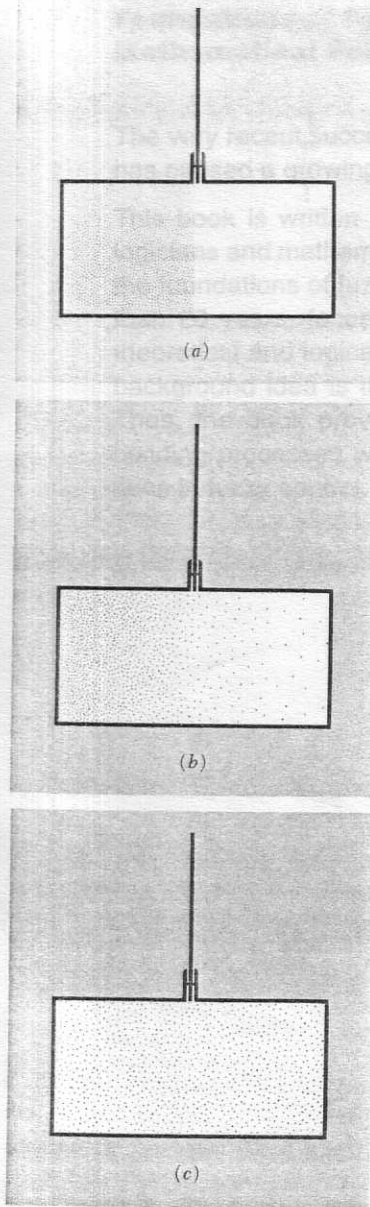
The important conclusion which we have reached in this section can thus be summarized as follows:

If an isolated system is in an appreciably nonrandom situation, it will (except for fluctuations which are unlikely to be large) change in time so as to approach ultimately its most random situation where it is in equilibrium.

(7)

Note that the preceding statement does *not* make any assertions about the *relaxation time*, i.e., the approximate time required for the system to reach its final equilibrium situation. The actual magnitude of this time depends sensitively on details of the system under consideration; it might be of the order of microseconds or of the order of centuries.

Fig. 1.11 The box of Fig. 1.10 is here shown (a) immediately after the partition has been removed, (b) a short time thereafter, and (c) a long time thereafter. A movie played backward would show the pictures in the reverse order (c), (b), (a).



Example

Referring to Fig. 1.10, consider again the box divided into two equal parts by a partition. The left half of the box contains N molecules of gas, while its right half is empty. Imagine now that the partition is suddenly removed—but only partly (as shown in Fig. 1.13) rather than fully (as in the previous experiment of Fig. 1.10). In both experiments the nonrandom situation immediately after the partition is removed (when the number n of molecules in the left half is equal to N) tends to change in time until the molecules become essentially uniformly distributed throughout the entire box (so that $n \approx \frac{1}{2}N$). But the time required until the final equilibrium condition is reached will be longer in the experiment of Fig. 1.13 than in that of Fig. 1.10.

Irreversibility

The statement (7) asserts that, when an isolated macroscopic system changes in time, it tends to do so in a very definite direction—namely, from a less random to a more random situation. We could observe the process of change by taking a movie of the system. Suppose now that we played the movie backward through a projector (i.e., that we ran the film through the projector in a direction opposite to that used in the original movie camera photographing the process). We would then observe on the screen the *time-reversed process*, i.e., the process that would occur if one imagined the direction of time to be reversed. The movie on the screen would look very peculiar indeed since it would portray a process in which the system changes from a more random to a much less random situation—something which one would almost never observe in actuality. Just by watching the movie on the screen, we could conclude with almost complete certainty that the movie is being played backward through the projector.

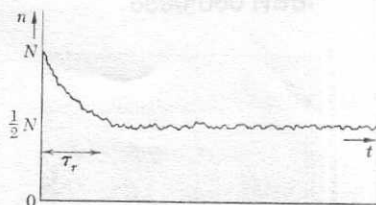


Fig. 1.12 Schematic illustration showing how the number n of molecules in the left half of the box in Fig. 1.11 varies as a function of the time t , beginning with the instant immediately after the partition is removed. The relaxation time is indicated by τ_r .

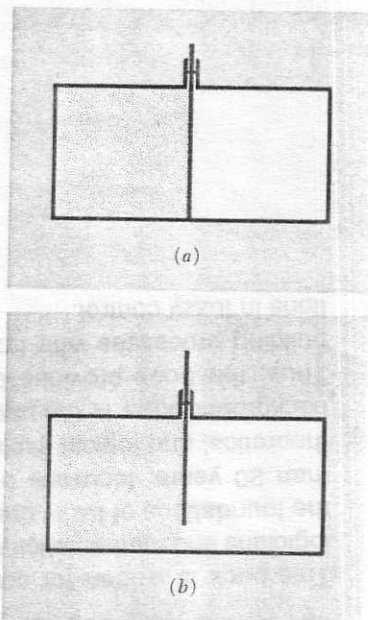


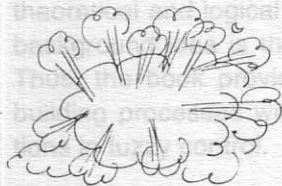
Fig. 1.13 The partition in (a) is suddenly removed partway only, as shown in (b).

The very recent success with even consumer products involving fuzzy tools has caused a growing interest in the field of fuzzy sets and fuzzy control.

Example

For example, suppose one filmed the process that takes place after the partition in Fig. 1.10 is suddenly removed. The movie played forward through the projector would show the gas spreading out (as indicated in Fig. 1.11) until it becomes essentially uniformly distributed throughout the entire box. Such a process is quite familiar. On the other hand, the movie played backward would show the gas, initially distributed uniformly throughout the box, concentrating itself spontaneously in the left half of the box so as to leave the right half empty. Such

a process is virtually never observed in actuality. This does not mean that this process is impossible, only that it is exceedingly improbable. What the backward movie shows *could* happen in actuality if all the molecules moved in an extremely special way.[†] But it is *exceedingly* improbable that all the molecules would ever move in this special way; indeed, it is as improbable as the occurrence of a fluctuation where $n = N$ (in a situation where the gas is in equilibrium throughout the entire box).



THE SATURDAY EVENING POST

Fig. 1.14 This cartoon is humorous because it portrays the reverse of an irreversible process. The indicated sequence of events *could* happen, but it is exceedingly unlikely that it ever would. (Reprinted by special permission of *The Saturday Evening Post* and James Frankfort, © 1965 The Curtis Publishing Company.)

A process is said to be *irreversible* if the time-reversed process (the one which would be observed in a movie played backward) is such that it would almost never occur in actuality. But all macroscopic systems not in equilibrium tend to approach equilibrium, i.e., situations of greatest randomness. Hence we see that all such systems exhibit irreversible behavior. Since in everyday life we are constantly surrounded by systems which are not in equilibrium, it becomes clear why time seems to have an unambiguous direction which allows us to distinguish clearly the past from the future. Thus we expect people to be born, grow up, and die. We never see the time-reversed process (in principle possible, but fantastically unlikely) where someone rises from his grave, grows progressively younger, and disappears into his mother's womb.

Note that there is nothing intrinsic in the laws of motion of the particles of a system which gives time a preferred direction. Indeed, suppose that one took a movie of the isolated gas in equilibrium, as shown in Fig. 1.4 (or considered the time dependence of the number n of molecules in one half of the box, as indicated in Fig. 1.5). Looking at this movie projected on a screen, we would have no way of

[†] Indeed, consider the molecules at some time t_1 after they have become uniformly distributed throughout the box. Now suppose that, at some subsequent time t_2 , each molecule would again be in exactly the same position as at the time t_1 and would have a velocity of equal magnitude but precisely opposite direction. Then each molecule would retrace its path in time. The gas would thus reconcentrate itself in the left half of the box.

telling whether it was running through the projector forward or backward. The preferred direction of time arises only when one deals with an isolated macroscopic system which is somehow known to be in a very special nonrandom situation at a specified time t_1 . If the system has been left undisturbed for a very long time and got into this situation as a result of a very rare spontaneous fluctuation in equilibrium, there is indeed nothing special about the direction of time. As already pointed out in connection with the peak X in Fig. 1.8, the system then tends to change toward the most random situation as time goes either forward or backward (i.e., both when the movie is played forward, and when it is played backward starting at the time t_1). The only other way the system can have got into the special nonrandom situation at the time t_1 is by virtue of interaction with some other system at some prior time. But in this case a specific direction of time is singled out by the knowledge that, before being left undisturbed, the system was made to interact with some other system at some time before the time t_1 .

Finally it is worth pointing out that the irreversibility of spontaneously occurring processes is a matter of degree. The irreversibility becomes more pronounced to the extent that the system contains many particles since the occurrence of an orderly situation then becomes increasingly unlikely compared to the occurrence of a random one.

Example

Consider a box which contains only a single molecule moving around and colliding elastically with the walls. If one took a movie of this system and then watched its projection on a screen, there would never be any way of telling whether the movie is being played forward or backward through the projector.

Consider now a box which contains N molecules of an ideal gas. Suppose that a movie of this gas is projected on a screen and portrays a process in which the molecules of the gas, originally distributed uniformly throughout the box, all become concentrated in the left half of the box. What could we conclude? If $N = 4$, this kind of process could, in actuality, occur relatively frequently as a result of a spontaneous fluctuation. (On the average, 1 out of every 16 frames of the film would show all the molecules in the left half of the box.) Hence we can-

not really tell with appreciable certainty whether the film is being played forward or backward. (See Fig. 1.15.) But if $N = 40$, this kind of process would very rarely occur in actuality as a result of a spontaneous fluctuation. (On the average, only 1 out of $2^N = 2^{40} \approx 10^{12}$ frames of the film would show all the molecules in the left half of the box.) It is much more likely that the film is being played backward and portrays the result of a prior intervention, e.g., the removal of a partition which had previously confined the molecules in the left half of the box. (See Fig. 1.17.) For an ordinary gas where $N \sim 10^{20}$, spontaneous fluctuations of the kind seen on the screen would almost never occur in actuality. One could then be almost completely certain that the film is being played backward.

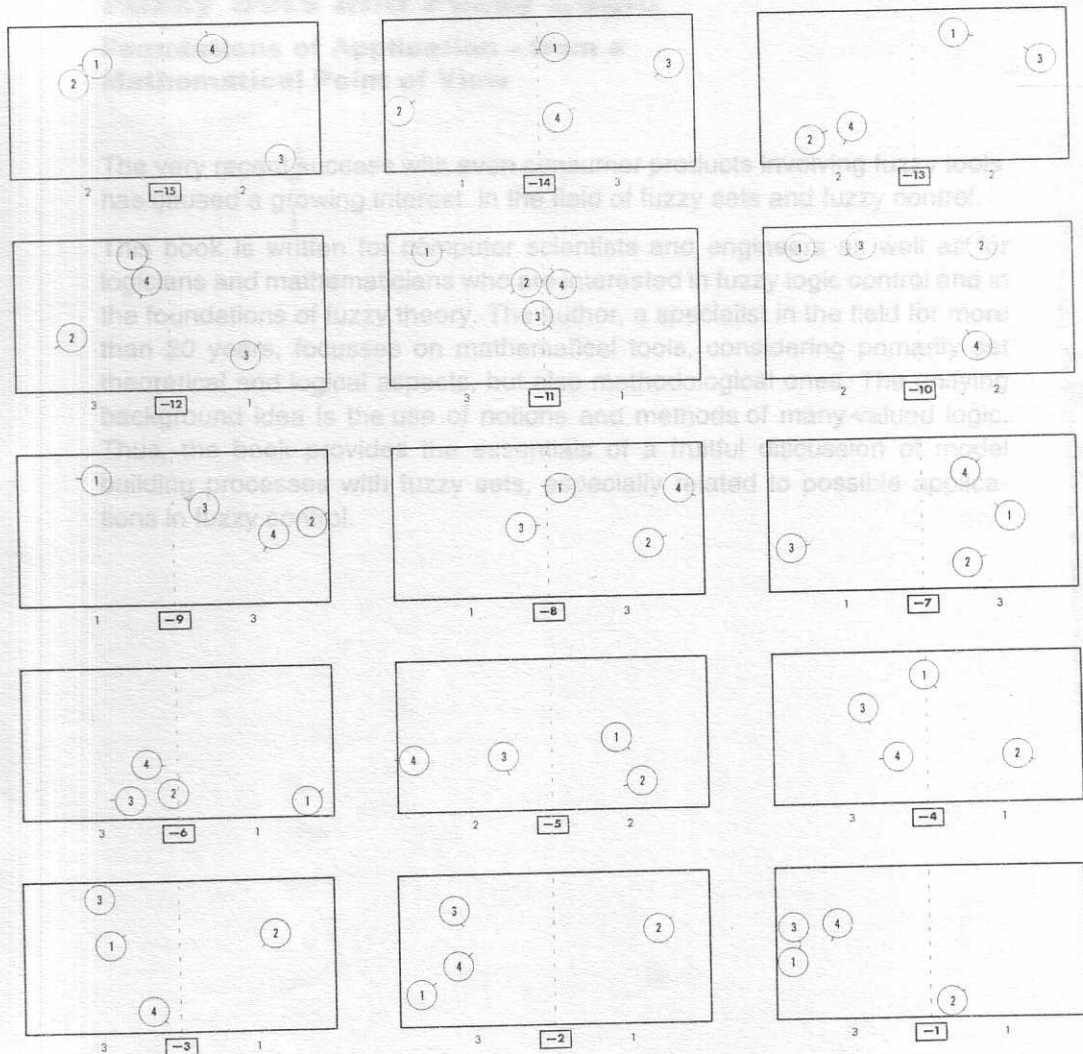


Fig. 1.15 Construction of a possible past history for Fig. 1.16. The pictures on this left-hand page were computed by starting with all the particles in the left half of the box in the positions shown in the frame $j = 0$ on the right-hand page and reversing the directions of all the assumed particle velocities in this frame. The resulting evolution of the system in time is then shown by the sequence of frames read in the order $j = 0, -1, -2, \dots, -15$. The short line segment emanating from each particle indicates the direction of its velocity in this case.

If the velocity of every particle on this left page is now imagined to be reversed in direction, then the sequence of frames in the order $j = -15, -14, \dots, -1, 0, 1, \dots, 14$ (extending over both left and right pages) represents a possible motion of the particles in time. This motion, starting from the very special initial situation prevailing in the frame $j = -15$ by virtue of the way it was constructed, then results in a fluctuation where all the particles are found in the left half of the box in the frame $j = 0$.

(w.r to the box in t=0 + this: Imaginary events)

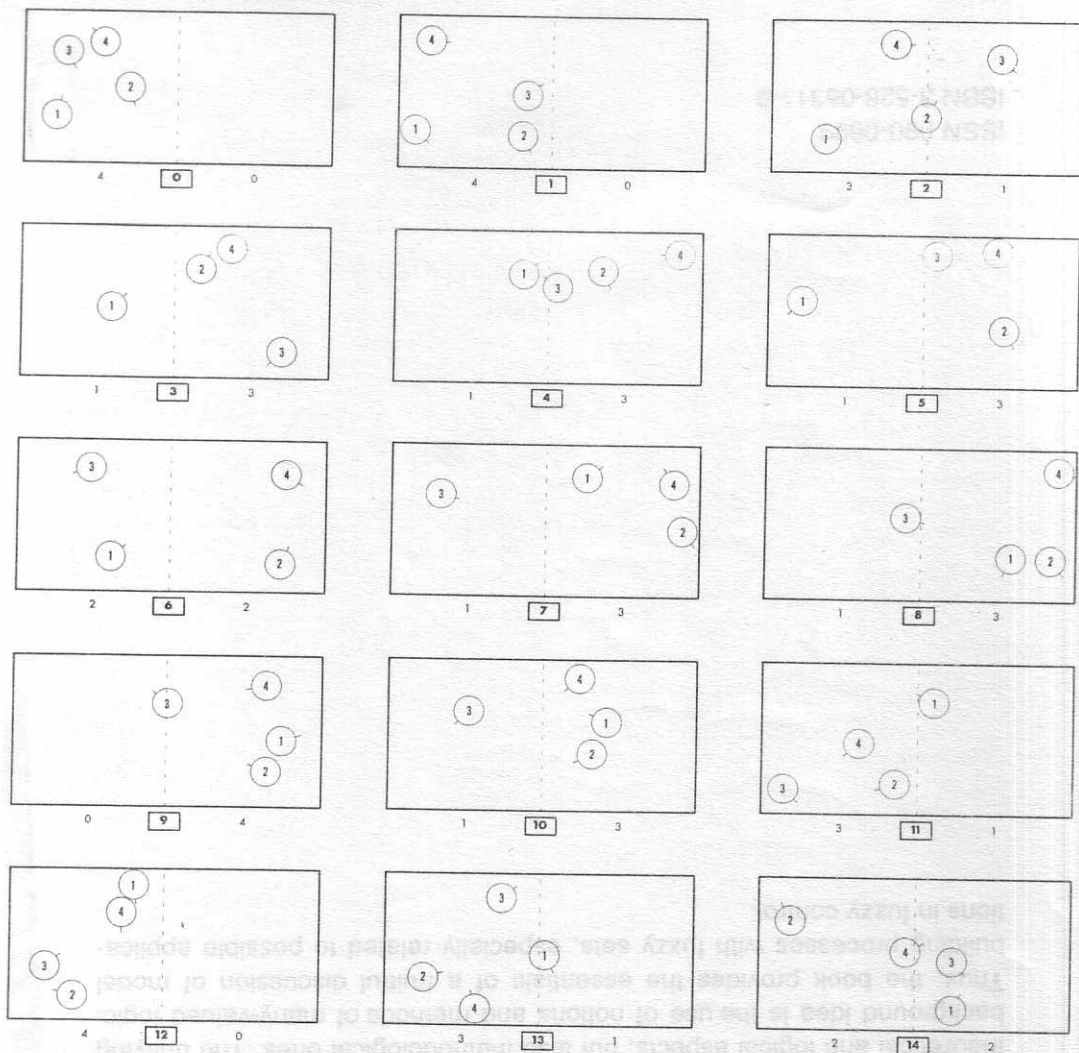


Fig. 1.16 Computer-made pictures showing 4 particles in a box. The pictures were constructed by starting with the special situation where all the particles are in the left half of the box in the positions shown in the frame $j = 0$ and are given some arbitrary assumed velocities. The resulting evolution of the system in

time is then shown by the sequence of frames $j = 0, 1, 2, \dots, 14$. The number of particles located in each half of the box is printed directly beneath that half. The short line segment emanating from each particle indicates the direction of the particle's velocity.

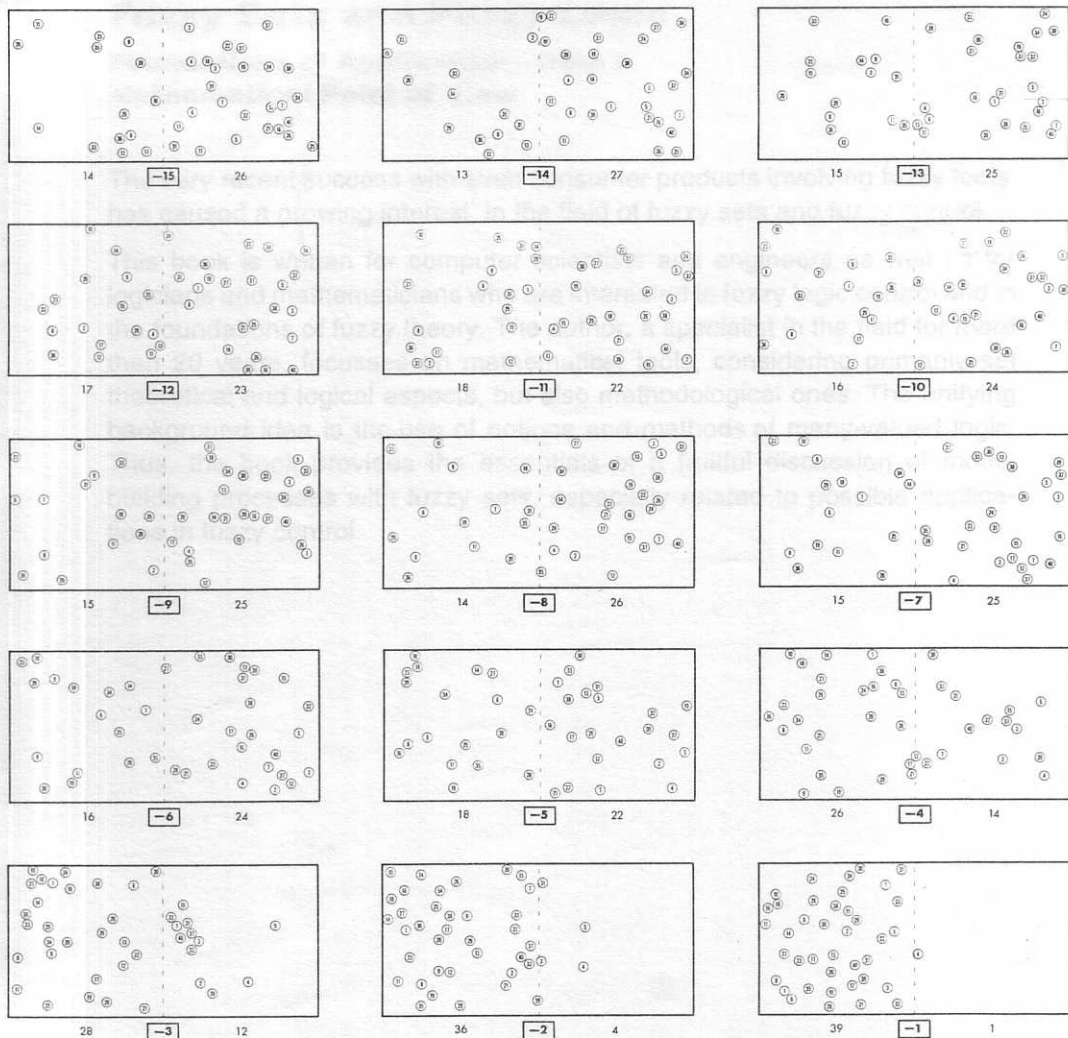


Fig. 1.17 Construction of a possible past history for Fig. 1.18. The pictures on this left-hand page were computed by starting with all the particles in the left half of the box in the positions shown in the frame $j = 0$ on the right-hand page and *reversing* the directions of all the assumed particle velocities in this frame. The resulting evolution of the system in time is then shown by the sequence of frames read in the order $j = 0, -1, -2, \dots, -15$. No velocities are indicated.

If the velocity of every particle on this left page is now imagined to be reversed in direction, then the sequence of frames in the order $j = -15, -14, \dots, -1, 0, 1, \dots, 14$ (extending over both left and right pages) represents a possible motion of the particles in time. This motion, starting from the very special initial situation prevailing in the frame $j = -15$ by virtue of the way it was constructed, then results in a fluctuation where all the particles are found in the left half of the box in the frame $j = 0$.

(Imaginary events: w.r. to the box, in $t < 0$)

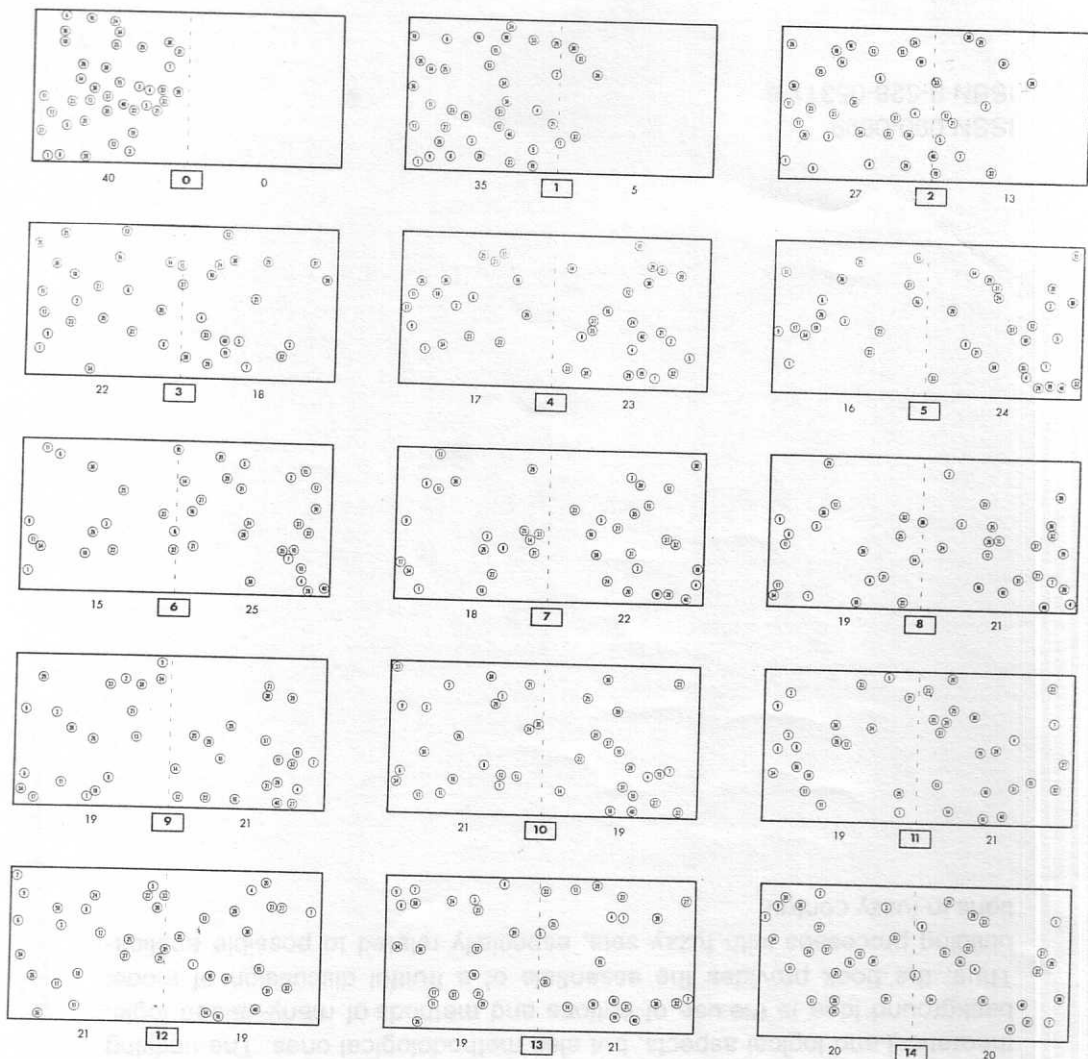


Fig. 1.18 Computer-made pictures showing 40 particles in a box. The pictures were constructed by starting with the special situation where all the particles are in the left half of the box in the positions shown in the frame $j = 0$ and are given some arbitrary assumed velocities.

The resulting evolution of the system in time is then shown by the sequence of frames $j = 0, 1, 2, \dots, 14$. The number of particles located in each half of the box is printed directly beneath that half. No velocities are indicated.

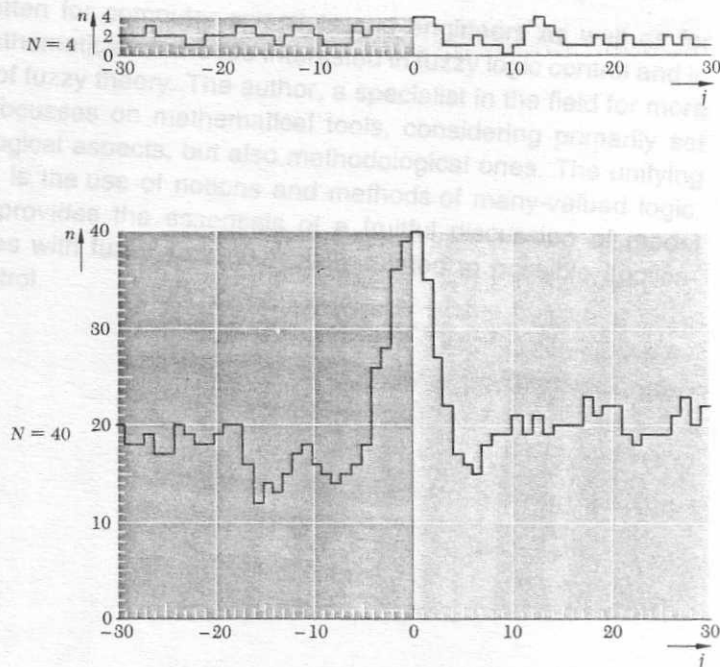


Fig. 1.19 The number n of particles in the left half of the box as a function of the frame index j or the time $t = j\tau_0$.^{*} The number n in the j th frame is indicated by a horizontal line extending from j to $j + 1$. The graphs describe Figs. 1.15 and 1.16 for $N = 4$ particles, and Figs. 1.17 and 1.18 for $N = 40$ particles, but contain information about more frames than were shown there. The right half of each graph shows the approach of the system to equilibrium. The entire domain of each graph shows the occurrence of a rare fluctuation which might occur in an equilibrium situation.

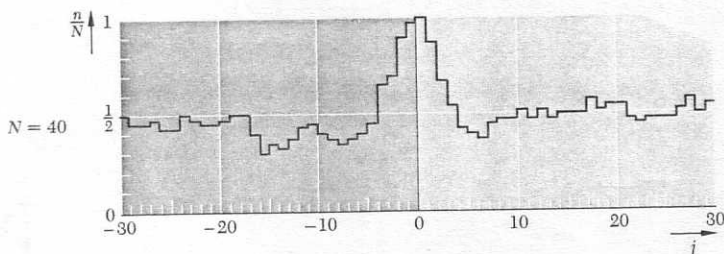
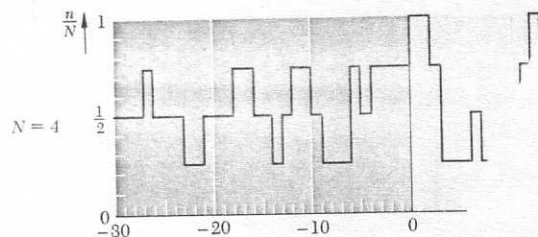


Fig. 1.20 The relative number n/N of particles in the left half of the box as a function of the frame index j or the time $t = j\tau_0$. The information presented is otherwise the same as that in Fig. 1.19.

1.3 Further Illustrations

By thinking in detail about the simple case of the ideal gas of N molecules, we have grappled with all the essential problems involved in understanding systems consisting of very many particles. Indeed, most of the remainder of this book will consist merely of the systematic elaboration and refinement of the ideas which we have already discussed. To begin, let us illustrate the universal applicability of the basic concepts which we have introduced by considering briefly a few further examples of simple macroscopic systems.

i) *Ideal system of N spins*

Consider a system of N particles each of which has a "spin $\frac{1}{2}$ " and an associated magnetic moment of magnitude μ_0 . The particles might be electrons, atoms having one unpaired electron, or nuclei such as protons. The concept of spin must be described in terms of quantum ideas. The statement that a particle has spin $\frac{1}{2}$ thus implies that a measurement of the component (along a specified direction) of the spin angular momentum of the particle can have only two discrete possible outcomes: The measured component can be either $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$ (where $\hbar$ denotes Planck's constant divided by 2π), i.e., the spin can be said to point either parallel or antiparallel to the specified direction. Correspondingly, the component (along the specified direction) of the magnetic moment of the particle can be either $+\mu_0$ or $-\mu_0$, i.e., the magnetic moment can also be said to point either parallel or antiparallel to the specified direction. For the sake of simplicity, we shall designate these two possible orientations as "up" or "down," respectively.†

The system of N particles with spin $\frac{1}{2}$ is thus quite similar to a collection of N bar magnets, each having a magnetic moment μ_0 which can point either up or down. For the sake of simplicity, we may regard the particles as essentially fixed in position, as they would be if they were atoms located at the lattice sites of a solid.‡ We shall call the system of spins *ideal* if the interaction between the spins is almost negligible. (This is the case if the average distance between the particles with spin is so large that the magnetic field produced by one moment at the position of another moment is small enough to be nearly negligible.)

† The magnetic moment of a particle may be antiparallel to its spin angular momentum. (This is the usual situation if the particle is negatively charged.) In this case the magnetic moment points down when the spin points up, and vice versa.

‡ If the particles are free to move in space, their translational motion can usually be treated separately from the orientation of their spins.

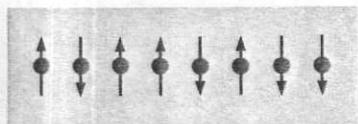
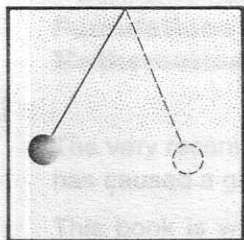


Fig. 1.21 A simple system of particles, each having spin $\frac{1}{2}$. Each spin can point either up or down.

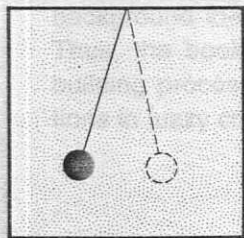
The ideal system of N spins has been described entirely in terms of quantum mechanics, but is otherwise completely analogous to the ideal gas of N molecules. In the case of the gas, each molecule moves about and collides occasionally with other molecules; hence it is found sometimes in the left and sometimes in the right half of the box. In the case of the system of spins, each magnetic moment interacts slightly with the other magnetic moments, so that its orientation occasionally changes; hence each magnetic moment is found to point sometimes up and sometimes down. In the case of the isolated ideal gas in equilibrium, each molecule is equally likely to be found in either the left or right half of the box. Similarly, in the case of the isolated system of spins in equilibrium in the absence of any externally applied magnetic field, each magnetic moment is equally likely to be found pointing either up or down. We can denote by n the number of spins pointing up, and by n' the number of spins pointing down. In equilibrium the most random situation where $n \approx n' \approx \frac{1}{2}N$ thus occurs most frequently, while fluctuations where n differs appreciably from $\frac{1}{2}N$ occur very rarely. Indeed, when N is large, nonrandom situations where n differs appreciably from $\frac{1}{2}N$ occur almost always as a result of prior interaction of the isolated system of spins with some other system.

(2) *Distribution of energy in an ideal gas*

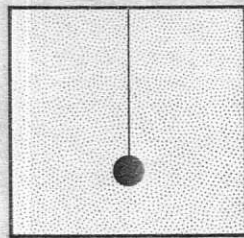
Consider again the isolated ideal gas of N molecules. We reached the general conclusion that the time-independent equilibrium situation attained by the system after a sufficiently long time corresponds to the most random distribution of the molecules. In our previous discussion we focused attention solely on the positions of the molecules. We then saw that the equilibrium of the gas corresponds to the most random distribution of the molecules in space, i.e., to the essentially uniform distribution of the molecules throughout the fixed volume of the box. But what can we say about the velocities of the molecules? Here it is useful to recall the fundamental mechanical principle that the total energy E of the gas must remain constant since the gas is an isolated system. This total energy E is equal to the sum of the energies of the individual gas molecules, since the potential energy of interaction between molecules is negligible. Hence the basic question becomes: How is the fixed total energy of the gas distributed over the individual molecules? (If a molecule is monatomic, its energy ϵ is, of course, merely its kinetic energy $\epsilon = \frac{1}{2}mv^2$, where m is its mass and v its velocity.) It is possible that one group of molecules might have very high energies while another group might have very low energies. But



(a)



(b)



(c)

Fig. 1.22 Oscillation of a pendulum immersed in a gas. The pendulum is shown successively (a) soon after it is set into oscillation, (b) a short time thereafter, and (c) after a very long time. A movie played backward would show the pictures in the reverse order (c), (b), (a).

this kind of situation is quite special and would not persist as the molecules collide with each other and thus exchange energy. The time-independent equilibrium situation which is ultimately attained corresponds therefore to the most random distribution of the total energy of the gas over all the molecules. Each molecule has then, on the average, the same energy and thus also the same speed.† In addition, since there is no preferred direction in space, the most random situation of the gas is that where the velocity of each molecule is equally likely to point in any direction.

Pendulum oscillating in a gas

Consider a pendulum which is set oscillating in a box containing an ideal gas. If the gas were not present, the pendulum would continue oscillating indefinitely with no change in amplitude. (We neglect frictional effects that might arise at the point of support of the pendulum.) But in the presence of the gas, the situation is quite different. The molecules of the gas collide constantly with the pendulum bob. In each such collision energy is transferred from the pendulum bob to a molecule, or vice versa. What is the net effect of these collisions? Again this question can be answered by our familiar general argument without the need to consider the collisions in detail.‡ The energy E_b (kinetic plus potential) of the pendulum bob plus the total energy E_g of all the gas molecules must remain constant since the total system is isolated (if one includes the earth which provides the gravitational attraction). If the energy of the bob were transferred to the gas molecules, it could be distributed over these many molecules in many different ways instead of remaining entirely associated with the bob. A much more random situation of the system would thus result. Since an isolated system tends to approach its most random situation, the pendulum gradually transfers practically all its energy to the gas molecules and thus oscillates with ever-decreasing amplitude. This is again a typical irreversible process. After the final equilibrium situation has been reached, the pendulum hangs vertically, except for very small oscillations about this position.

† This does *not* mean that each molecule has the same energy at any one time. The energy of any one molecule fluctuates quite appreciably in the course of time as a result of its collisions with other molecules. But when each molecule is observed over a sufficiently long time interval τ , its average energy over that time interval is the same as that of any other molecule.

‡ A detailed analysis would argue that the pendulum bob suffers more collisions per unit time with the molecules located on the side toward which the bob moves than with the molecules located on the other side. As a result, collisions in which the bob loses energy to a molecule are more frequent than collisions in which it gains energy from a molecule.

Note one further point of interest. In the initial nonrandom situation, where the pendulum bob has a large amount of energy associated with it, this energy can be exploited to do useful work on a macroscopic scale. For example, the pendulum bob could be made to hit a nail so as to drive it some distance into a piece of wood. After the final equilibrium has been reached, the energy of the pendulum bob has not been lost; it has merely become redistributed over the many molecules of the gas. But there is now no easy way to use this energy to do the work necessary to drive the nail into the wood. Indeed, this would necessitate some method for concentrating energy, randomly distributed over the many gas molecules moving in many directions, so that it can exert a net force in only one particular direction over an appreciable distance.

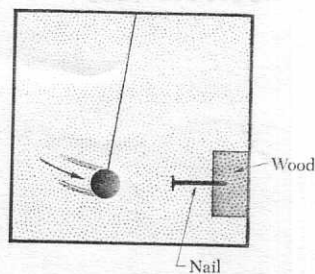


Fig. 1.23 Arrangement in which the pendulum bob can be made to strike a nail and thus do work on it by driving it into a piece of wood.

1.4 Properties of the Equilibrium Situation

(1) Simplicity of the equilibrium situation

The discussion of the preceding sections shows that the equilibrium situation of a macroscopic system is particularly simple. The reasons are the following:

(i) The macrostate of a system in equilibrium is time-independent, except for ever-present fluctuations. Quite generally, the macrostate of a system can be described by certain *macroscopic parameters*, i.e., parameters which characterize the properties of the system on a large scale. (For example, the number n of molecules located in the left half of a box of gas is such a macroscopic parameter.) When the system is in equilibrium, the average values of all its macroscopic parameters remain constant in time, although the parameters themselves may exhibit fluctuations (ordinarily quite small) about their average values. The equilibrium situation of a system is, therefore, simpler to treat than the more general nonequilibrium case where some macroscopic parameters of the system tend to change in time.

(ii) The macrostate of a system in equilibrium is, except for fluctuations, the most random macrostate of the system under the specified conditions. The system in equilibrium is thus characterized in a unique way. In particular, this has the following implications:

(a) The equilibrium macrostate of a system is independent of its past history. For example, consider an isolated gas of N molecules in a box. These molecules may originally have been confined by a partition to one half of the box or to one quarter of the box (the total energy of the molecules being assumed the same in each case). But after the partition is removed and equilibrium has been attained, the macrostate of the gas is the same in both cases; it corresponds merely to the uniform distribution of all the molecules through the entire box.

(b) The equilibrium macrostate of a system can be completely specified by very few macroscopic parameters. For example, consider again the isolated gas of N identical molecules in a box. Suppose that the volume of the box is V , while the constant total energy of all the molecules is E . If the gas is in equilibrium and thus known to be in its most random situation, then the molecules must be uniformly distributed throughout the volume V and must, on the average, share equally the total energy E available to them. A knowledge of the macroscopic parameters V and E is, therefore, sufficient to conclude that the average number $\bar{n}_s$ of molecules in any subvolume V_s of the box is $\bar{n}_s = N(V_s/V)$, and that the average energy $\bar{\epsilon}$ per molecule is $\bar{\epsilon} = E/N$. If the gas were not in equilibrium, the situation would of course be much more complicated. The distribution of molecules would ordinarily be highly nonuniform and a mere knowledge of the total number N of molecules in the box would thus be completely insufficient for determining the average number $\bar{n}_s$ of molecules in any given subvolume V_s of the box.

(2) *Observability of fluctuations*

Consider a macroscopic parameter describing a system consisting of many particles. If the number of particles in the system is large, the relative magnitude of the fluctuations exhibited by the parameter is ordinarily very small. Indeed, it is often so small as to be utterly negligible compared to the average value of the parameter. As a result, we remain usually unaware of the existence of fluctuations when we are dealing with large macroscopic systems. On the other hand, the ever-present fluctuations may be readily observed and may become of great practical importance if the macroscopic system under consideration is fairly small or if our methods of observation are quite sensitive. Several examples will serve to illustrate these comments.

(3) Density fluctuations in a gas

Consider an ideal gas which is in equilibrium and consists of a large number N of molecules confined within a box of volume V . Focus attention on the number n_s of molecules located within some specified subvolume V_s inside the box. This number n_s fluctuates in time about an average value

$$\bar{n}_s = \frac{V_s}{V} N,$$

the magnitude of its fluctuation at any time being given by the difference

$$\Delta n_s = n_s - \bar{n}_s.$$

If we consider the left half of the box as the part of interest, $V_s = \frac{1}{2}V$ and $\bar{n}_s = \frac{1}{2}N$. When V_s is large, the average number $\bar{n}_s$ of molecules is also large. In accordance with our discussion of Sec. 1.1, the only fluctuations which occur with appreciable frequency are then those sufficiently small so that $|\Delta n_s| \ll \bar{n}_s$.

On the other hand, suppose that we wanted to investigate the scattering of light by a substance. Then we would be interested in knowing what happens in a

volume element V_s having linear dimensions of the order of the wavelength of light. [Since the wavelength of visible light (about 5×10^{-5} cm) is much larger than atomic dimensions, such a volume element is still macroscopic in size, although it is small.] If the number of molecules in every such volume element were the same (as is very nearly the case in a solid such as glass), then the substance under consideration would be spatially uniform and would merely refract a beam of light without scattering it. But in the case of our ideal gas, the average number $\bar{n}_s$ of molecules in a volume as small as V_s is quite small and fluctuations Δn_s in the number n_s of molecules within V_s are no longer negligible compared to $\bar{n}_s$. The gas can, therefore, be expected to scatter light to an appreciable extent. Indeed, the fact that the sky does not look black is due to the fact that light from the sun is scattered by the gas molecules of the atmosphere. The blue color of the sky thus provides visible evidence for the importance of fluctuations.

(3) Fluctuations of a torsion pendulum

Consider a thin fiber stretched between two supports (or suspended from one support under the influence of gravity) and carrying an attached mirror. When the mirror turns through a small angle, the twisted fiber provides a restoring torque. The mirror is thus capable of performing small angular oscillations and constitutes accordingly a torsion pendulum. Since the restoring torque of a thin fiber can be made very small and since a beam of light reflected from the mirror provides a very effective way of detecting small angular deflections of the mirror, a torsion fiber is commonly used for very sensitive measurements of small torques. For example, it may be recalled

that a torsion pendulum was used by Cavendish to measure the universal constant of gravitation and by Coulomb to measure the electrostatic force between charged bodies.

When a sensitive torsion pendulum is in equilibrium, its mirror is not perfectly still, but can be seen to perform erratic angular oscillations about its average equilibrium orientation. (The situation is analogous to that of the ordinary pendulum, discussed in Sec. 1.3, which exhibits in equilibrium small fluctuations about its vertical position.) These fluctuations are ordinarily caused by the random impacts of the surrounding air molecules on the mirror.

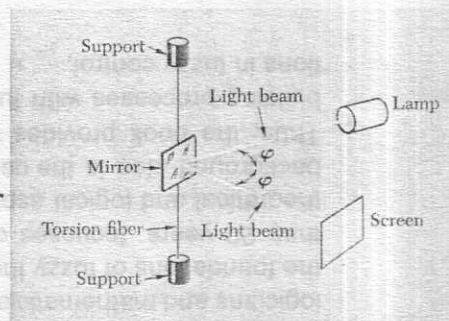


Fig. 1.24 Torsion pendulum formed by a mirror mounted on a thin fiber. A beam of light reflected from the mirror indicates the angle of rotation ϕ of the mirror.

[The fluctuations of the mirror would be modified, but would not disappear, even if all the molecules of the surrounding gas were removed. In that case the total energy of the torsion pendulum would still consist of two parts, the energy E_ω due to the angular velocity of the mirror moving as a whole, plus the energy E_i due to the internal motion of all the atoms of the mirror and fiber. (The atoms are free to perform small vibra-

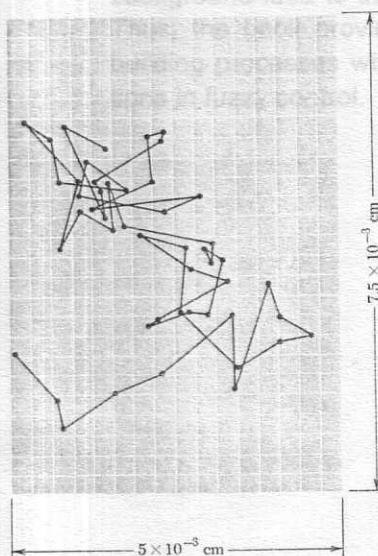
tions about their sites in the solids which constitute the mirror and fiber.) Although the total energy $E_\omega + E_i$ of the torsion pendulum is constant, fluctuations do occur in the way this energy is shared between E_ω and E_i . Any fluctuation in which E_ω gains energy at the expense of the internal motion of the atoms thus results in an increased angular velocity of the mirror, and vice versa.]

Brownian motion of a particle

Small solid particles, about 10^{-4} cm in size, can be introduced into a drop of liquid and then observed under a microscope. Such particles are not at rest, but are seen to be constantly moving about in a highly irregular way. This phenomenon is called *Brownian motion* because it was first observed in the last century by an English botanist named Brown. He did not understand the origin of the phenomenon. It remained for Einstein in 1905 to give the correct explanation in terms of random fluctuations to

be expected in equilibrium. A solid particle is subject to a fluctuating net force due to the many random collisions of the particle with the molecules of the liquid. Since the particle is small, the number of molecules with which it collides per unit time is relatively small and accordingly fluctuates appreciably. In addition, the mass of the particle is so small that any collision has a noticeable effect on the particle. The resulting random motion of the particle thus becomes large enough to be observable.

Fig. 1.25 Brownian motion of a solid particle, 10^{-4} cm in diameter, suspended in water and observed through a microscope. The three-dimensional motion of such a particle, as seen projected in the horizontal plane of the field of view of the microscope, is shown by this diagram where the lines join consecutive positions of the particle as observed at 30-second intervals. [The data are from J. Perrin, *Atoms*, p. 115 (D. Van Nostrand Company, Inc., Princeton, N.J., 1916).]



Voltage fluctuations across a resistor

If an electrical resistor is connected across the input terminals of a sensitive electronic amplifier, the output of the amplifier is observed to exhibit random voltage fluctuations. Neglecting noise originating in the amplifier itself, the basic reason is the random Brownian motion of the electrons in the resistor. Suppose, for example, that this random

motion leads to a fluctuation where the number of electrons is greater in one half of the resistor than in the other. The resulting charge difference then leads to an electric field in the resistor, and hence to a potential difference across its ends. Variations in this potential difference thus give rise to the voltage fluctuations amplified by the electronic instrument.

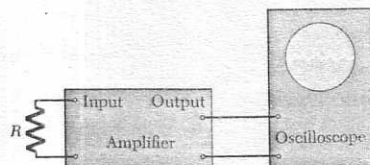


Fig. 1.26 A resistor R connected to the input terminals of a sensitive amplifier whose output is displayed on an oscilloscope.

† The existence of fluctuations can have important practical consequences. This is particularly true whenever one is interested in measuring small effects or signals, since these may be obscured by the intrinsic fluctuations ever-present in the measuring instrument. (These fluctuations are then said to constitute *noise* since their presence makes measurements difficult.) For example, it is difficult to use a torsion fiber to measure a torque which is so small that the angular deflection produced by it is less than the magnitude of the intrinsic fluctuations in angular position displayed by the mirror. Similarly, in the case of the resistor connected to the amplifier, it is difficult to measure an applied voltage across this resistor if this voltage is smaller than the magnitude of the intrinsic voltage fluctuations always present across the resistor.†

1.5 Heat and Temperature

Macroscopic systems which are not isolated can interact and thus exchange energy. One obvious way in which this can happen is if one system does macroscopically recognizable work upon some other system. For example, in Fig. 1.28 the compressed spring A' exerts a net force on the piston confining the gas A . Similarly, in Fig. 1.29, the compressed gas A' exerts a net force on the piston confining the gas A . When the piston is released and moves through some macroscopic distance, the force exerted by A' performs some work† on the system A .

It is, however, quite possible that two macroscopic systems can interact under circumstances where no macroscopic work is done. This type of interaction, which we shall call *thermal interaction*, occurs because energy can be transferred from one system to the other system on an atomic scale. The energy thus transferred is called *heat*. Imagine that the piston in Fig. 1.29 is clamped in position so that it cannot move. In this case one system cannot do macroscopic work upon the other, irrespective of the net force exerted on the piston. On the other hand, the atoms of system A do interact (or collide) with each other almost constantly and thus exchange energy among themselves.§

† Typically, voltages of the order of one microvolt or less may become difficult to measure. By averaging the measurement over a sufficiently long time, one may, however, discriminate against the random fluctuations in favor of the applied signal which does not fluctuate in time.

‡ The term *work* is used here in its usual sense familiar from mechanics and is thus defined basically as a force multiplied by the distance through which it acts.

§ If each molecule of the gas consists of more than one atom, different molecules can exchange energy by colliding with each other; furthermore, the energy of a single molecule can also be exchanged between its constituent atoms as a result of interaction between them.

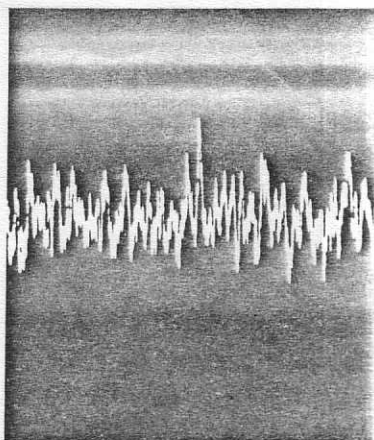


Fig. 1.27 Actual photograph of the noise output voltage displayed on an oscilloscope in the experimental arrangement of Fig. 1.26. (Photograph by courtesy of Dr. F. W. Wright, Jr., University of California, Berkeley.)

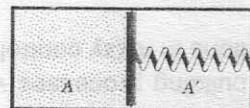


Fig. 1.28 The compressed spring A' does work on the gas A when the piston moves through some net macroscopic distance.



Fig. 1.29 The compressed gas A' does work on the gas A when the piston moves through some net macroscopic distance.

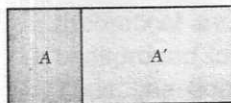


Fig. 1.30 Schematic diagram showing two general systems A and A' in thermal contact with each other.

Similarly, the atoms of system A' exchange energy among themselves. At the piston, which forms the boundary between the gases A and A' , the atoms of A interact with the atoms of the piston, which then interact among themselves and then in turn interact with the atoms of A' . Hence energy can be transferred from A to A' (or from A' to A) as a result of many successive interactions between the atoms of these systems.

Consider then any two systems A and A' in thermal interaction with each other. For example, the two systems might be the two gases A and A' just discussed; or A might be a copper block immersed in a system A' consisting of a container filled with water. Let us denote by E the energy of system A (i.e., the total energy, kinetic plus potential, of all the atoms in A); similarly let us denote by E' the energy of system A' . Since the combined system $A^* \equiv A + A'$ consisting of both A and A' is supposed to be insulated, the total energy

$$E + E' = \text{constant.}^\dagger \quad (8)$$

The question arises, however, how this total energy is actually distributed between the systems A and A' . In particular, suppose that the systems A and A' are in equilibrium with each other, i.e., that the combined system A^* is in equilibrium. Except for small fluctuations, this equilibrium situation of A^* must then correspond to the most random distribution of energy in this system.

Let us first discuss the simple situation where the systems A and A' are two ideal gases consisting of molecules of the same kind. [For example, both A and A' might consist of nitrogen (N_2) molecules.] In this case the most random situation of the combined system A^* is clearly that where its total energy $E + E'$ is shared indiscriminately among all the identical molecules of A^* . Each molecule of A and A' should then have the same average energy. In particular, the average energy $\bar{\epsilon}$ of a molecule of gas A should be the same as the average energy $\bar{\epsilon}'$ of a molecule of gas A' , i.e.,

$$\text{in equilibrium,} \quad \bar{\epsilon} = \bar{\epsilon}'. \quad (9)$$

Of course, if there are N molecules in gas A and N' molecules in gas A' , then

$$\bar{\epsilon} = \frac{E}{N} \quad \text{and} \quad \bar{\epsilon}' = \frac{E'}{N'}. \quad (10)$$

[†] In the case of the system of Fig. 1.29, we assume for the sake of simplicity that the container walls and the piston are so thin that their energies are negligibly small compared to the energies of the gases.

Hence the condition (9) can also be written in the form

$$\frac{E}{N} = \frac{E'}{N'}.$$

Suppose that the gases A and A' are initially separated from each other and separately in equilibrium. We denote their energies under these circumstances by E_i and E'_i , respectively. Imagine now that the systems A and A' are placed in contact with each other so that they are free to exchange energy by thermal interaction. There are two cases which may arise:

(i) Ordinarily, the initial energies E_i and E'_i of the systems are such that the average initial energy $\bar{\epsilon}_i = E_i/N$ of a molecule of A is not the same as the average initial energy $\bar{\epsilon}'_i = E'_i/N'$ of a molecule of A' ; i.e.,

$$\text{ordinarily,} \quad \bar{\epsilon}_i \neq \bar{\epsilon}'_i. \quad (11)$$

In this case the initial distribution of energy in the combined system A^* is quite nonrandom and will not persist in time. Instead, the systems A and A' will exchange energy until they ultimately attain the equilibrium situation corresponding to the most random distribution of energy, that where the average energy per molecule is the same in both systems. The energies E_f and E'_f of the systems A and A' in the final equilibrium situation must then be such that

$$\bar{\epsilon}_f = \bar{\epsilon}'_f \quad \text{or} \quad \frac{E_f}{N} = \frac{E'_f}{N'}. \quad (12)$$

In the interaction process leading to the final equilibrium situation the system with the smaller average initial energy per molecule thus gains energy, while the system with the larger average initial energy per molecule loses energy. Of course, the total energy of the isolated combined system A^* remains constant so that

$$E'_f + E_f = E'_i + E_i.$$

$$\text{Thus} \quad \Delta E + \Delta E' = 0, \quad (13)$$

$$\text{or} \quad Q + Q' = 0, \quad (14)$$

where we have written

$$\begin{aligned} Q &\equiv \Delta E \equiv E_f - E_i, \\ Q' &\equiv \Delta E' \equiv E'_f - E'_i. \end{aligned} \quad (15)$$

The quantity Q is called *the heat absorbed by A* in the interaction process and is defined as the increase in energy of A resulting from the thermal interaction process. A similar definition holds for the heat Q' absorbed by A' .

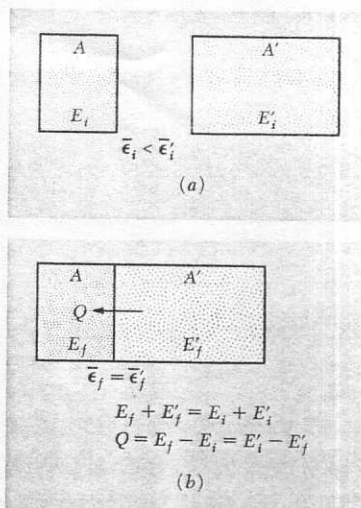


Fig. 1.31 Two gases A and A' , consisting of molecules of the same kind, are initially separated from each other in (a). They are then brought into thermal contact with each other in (b) and are allowed to exchange heat until they reach equilibrium. The energy of a gas is denoted by E , the average energy per molecule by $\bar{\epsilon}$.

Note that the heat $Q = \Delta E$ absorbed by a system may be either positive or negative. Indeed, in the thermal interaction between two systems, one loses energy while the other gains energy; i.e., in (14) either Q is positive and Q' is negative, or vice versa. By definition, the system which gains energy by absorbing a positive amount of heat is said to be the *colder* system; on the other hand, the system which loses energy by absorbing a negative amount of heat (i.e., by "giving off" a positive amount of heat) is said to be the *warmer* or *hotter* system.

(ii) A special case may arise where the initial energies E_i and E_i' of the systems A and A' are such that the average energy of a molecule of A happens to be the same as that of a molecule of A' ; i.e., it may happen that

$$\bar{\epsilon}_i = \bar{\epsilon}_i'. \quad (16)$$

In this case, when the systems A and A' are brought into thermal contact with each other, they automatically give rise to the most random situation of the combined system A^* . Thus the systems remain in equilibrium and there occurs no net transfer of energy (or heat) between the systems.

Temperature

Consider now the general case of thermal interaction between two systems A and A' . These systems might be different gases whose molecules have different masses or consist of different kinds of atoms; one or both of the systems might also be liquids or solids. Although the conservation of energy (8) is still valid, it now becomes more difficult to characterize the equilibrium situation corresponding to the most random distribution of the energy over all the atoms of the combined system $A + A'$. Qualitatively, most of our preceding comments concerning the case of two similar gases ought, however, still to be applicable. What we should then expect (and shall explicitly show later on) is the following: Each system, such as A , is characterized by a parameter T (conventionally called its "absolute temperature") which is related to the average energy per atom in the system. In the equilibrium situation corresponding to the most random distribution of energy we then expect, analogously to (9), a condition of the form

$$T = T'. \quad (17)$$

It is not possible to define the concept of absolute temperature more precisely until we have given a more precise definition of the concept of randomness (applicable to the distribution of energy over unlike atoms). It is, however, easy to introduce the concept of a temperature

(not the “absolute temperature”) measured with a particular thermometer. By a *thermometer* we mean any small macroscopic system M arranged so that only one of its macroscopic parameters changes when the system M absorbs or gives off heat. This parameter is called the *thermometric parameter* of the thermometer and will be denoted by the Greek letter θ . For example, the familiar mercury or alcohol thermometer is a special example of a thermometer. Here the length L of the column of liquid in the glass capillary tube of the thermometer changes when the energy of the liquid changes as a result of a transfer of heat. The thermometric parameter θ of this kind of thermometer is then the length L . If the thermometer M is placed in thermal contact with some other system A and allowed to come to equilibrium, its thermometric parameter θ assumes some value θ_A . This value θ_A is called the *temperature of the system A with respect to the particular thermometer M* .†

The usefulness of a thermometer is made apparent by the following considerations. Suppose that the thermometer M is placed in thermal contact first with a system A and then with a system B . In each case the thermometer is allowed to come to equilibrium; it then indicates the respective temperatures θ_A and θ_B . Two cases may arise: either $\theta_A \neq \theta_B$ or $\theta_A = \theta_B$. It is a familiar experimental fact (which will later be justified theoretically) that if $\theta_A \neq \theta_B$, the systems A and B will exchange heat when brought into thermal contact with each other; but, if $\theta_A = \theta_B$, the systems will not exchange heat on thermal contact. The temperature θ of a system measured with respect to the particular thermometer is thus a very useful parameter characterizing the system, since a knowledge of the temperature allows the following predictive statement: If two systems are brought into thermal contact with each other, they will exchange heat if their temperatures are unequal and will not exchange heat if their temperatures are equal.

1.6 Typical Magnitudes

The preceding sections have shown us in broad outline how the behavior of macroscopic systems can be understood in terms of their constituent molecules or atoms. Our considerations have, however, been quite qualitative. To complete our preliminary orientation, we should also like to acquire some perspective about representative orders of

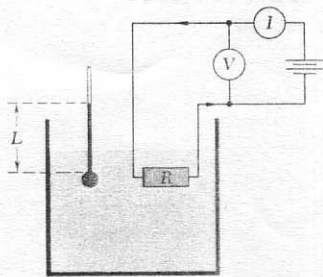


Fig. 1.32 Two different kinds of thermometers in thermal contact with a system consisting of a container filled with a liquid. One thermometer is a mercury thermometer; its thermometric parameter is the length L of the mercury column in the glass capillary tube. The other thermometer is an electrical resistor R made, for example, either of platinum wire or carbon; its thermometric parameter is its electrical resistance R (determined by passing a small current I through the resistor and measuring the voltage V across it).

† The system A is supposed to be much larger than the thermometer M so that it is disturbed negligibly by the small amount of energy gained or lost to the thermometer. Note also that, according to our definition, the temperature measured by a mercury thermometer is a length and is thus to be measured in units of centimeters.

Could we understand the behavior of microscopic systems (elementary particles, wave-particle duality, etc) provided their generators become understood.

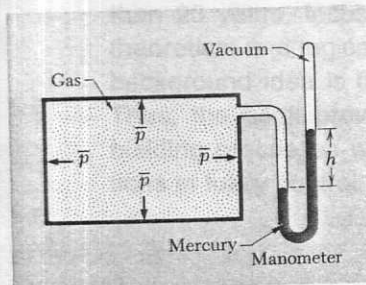


Fig. 1.33 A gas whose average pressure $\bar{p}$ is measured by means of a manometer consisting of a U-tube filled with mercury. In order to be in mechanical equilibrium, the mercury levels must adjust themselves so that the mercury column of height h has a weight, per unit cross-sectional area, equal to the pressure exerted by the gas.

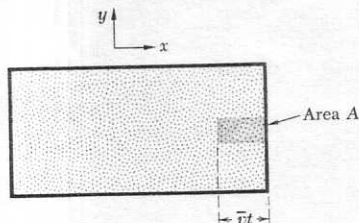


Fig. 1.34 Diagram illustrating collisions of gas molecules with an area A of a container wall. (The z axis points out of the paper.)

magnitude. For example, it would be worth knowing how fast a typical molecule moves or how frequently it collides with other molecules. To answer such questions, we may again turn to the simple case of an ideal gas.

Pressure of an ideal gas

When a gas is confined within a container, the many collisions of the gas molecules with the walls of the container give rise to a net force on each element of area of the walls. The force per unit area is called the *pressure* p of the gas. The average pressure $\bar{p}$ exerted by the gas is readily measured, for example by means of a manometer. The gas pressure should be calculable in terms of molecular quantities; conversely, it should then be possible to use the measured gas pressure to deduce the magnitudes of molecular quantities. We shall begin, therefore, by giving a simple approximate calculation of the pressure exerted by an ideal gas. m.g. $\rightarrow P$

To be specific, we shall consider an ideal gas of N molecules, each having a mass m . We suppose that the gas is in equilibrium and that it is confined within a box in the shape of a rectangular parallelepiped having a volume V . The number of molecules per unit volume is conveniently denoted† by $n \equiv N/V$. The edges of the box can be assumed to be parallel to a set of cartesian x , y , z axes, as shown in Fig. 1.34.

We focus attention on a wall of the box, e.g., the right wall which is perpendicular to the x axis. Let us first ask how many molecules collide with an area A of this wall during some short time interval t . All the molecules do not have the same velocity at any one time. But since we are satisfied with approximate results, we can simplify our considerations by assuming that each molecule moves with the same speed, equal to its average speed $\bar{v}$. The molecules move, however, in random directions so that, on the average, one-third of them (or $\frac{1}{3}n$ molecules per unit volume) move predominantly along the x axis, one-third along the y axis, and one-third along the z axis. Of the $\frac{1}{3}n$ molecules per unit volume moving predominantly along the x axis, one-half (or $\frac{1}{6}n$ molecules per unit volume) move in the $+x$ direction toward the area A , while the remaining half move in the $-x$ direction away from the area A . Any molecule having a velocity directed predominantly in the $+x$ direction travels during the short time t a distance $\bar{v}t$ in the $+x$ direction. If such a molecule is located within a distance

† The fact that the symbol n was previously used in a different context to denote the number of molecules in one half of a box of gas should not lead to confusion.

+ $\bar{v}t$ from the area A of the wall, it will thus strike this area in the time t ; but if it lies further than a distance $\bar{v}t$ from the area A , it will not reach this area and thus will not collide with it.† Hence the average number of molecules which collide with the area A during the time t is simply equal to the average number of molecules which have a velocity predominantly in the $+x$ direction and which are contained within the cylinder of area A and length $\bar{v}t$. It is, therefore, obtained by multiplying $\frac{1}{6}n$ (the average number of molecules per unit volume which have their velocity predominantly in the $+x$ direction) by the volume $A\bar{v}t$ of the cylinder and is thus given by

$$(\frac{1}{6}n)(A\bar{v}t).$$

• If we divide this result by the area A and the time t , we obtain an approximate expression for $\mathcal{F}_0$, the average number of molecules striking unit area of a wall per unit time (or the molecular *flux density*). Thus

$$\boxed{\mathcal{F}_0 \approx \frac{1}{6}n\bar{v}.} \quad (18)$$

• Let us now calculate the average force exerted by the colliding molecules on a unit area of the wall. When such a molecule, moving predominantly in the $+x$ direction, strikes the wall, its kinetic energy $\frac{1}{2}m\bar{v}^2$ remains unchanged. (This must be true, at least on the average, since the gas is in equilibrium.) The *magnitude* of the momentum of the molecule must then, on the average, also remain unchanged; i.e., the molecule, approaching the right wall with momentum $m\bar{v}$ in the $+x$ direction, must have a momentum $-m\bar{v}$ in this direction after it rebounds from the wall. The $+x$ component of momentum of the molecule changes then by an amount $-m\bar{v} - m\bar{v} = -2m\bar{v}$ as a result of the collision with the wall. Correspondingly it follows, by the principle of conservation of momentum, that the wall gains in the collision an amount of momentum $+2m\bar{v}$ in the $+x$ direction. But the average force exerted on the wall by the gas molecules is, by Newton's second law, equal to the average rate of change of momentum experienced by the wall as a result of the molecular collisions. The average force exerted on unit area of the wall (i.e., the average pressure $\bar{p}$ on the wall) is then simply obtained by multiplication, i.e.,

† Since the time interval t can be considered arbitrarily short, much shorter than the average time between collisions of the molecules among themselves, collisions of the given molecule with other molecules are very unlikely to occur during the time t and can be disregarded.

$$\bar{p} = \left[\begin{array}{l} \text{the average momentum} \\ 2m\bar{v} \text{ gained by the wall} \\ \text{in one molecular collision} \end{array} \right] \times \left[\begin{array}{l} \text{the average number of colli-} \\ \text{sions experienced per unit time} \\ \text{by a unit area of the wall} \end{array} \right].$$

$$\text{Thus} \quad \bar{p} \approx (2m\bar{v})\bar{f}_0 = (2m\bar{v})\left(\frac{1}{6}n\bar{v}\right)$$

$$\text{or} \quad \boxed{\bar{p} \approx \frac{1}{3}nm\bar{v}^2} \quad (19)$$

- As would be expected, the pressure $\bar{p}$ is seen to be increased (i) if n is made large so that there are more molecules colliding with the walls, and (ii) if $\bar{v}$ is increased so that the molecules collide with the walls more frequently and give up more momentum per collision.

Since the average kinetic energy $\epsilon^{(k)}$ of a molecule is approximately given by†

$$\overline{\epsilon^{(k)}} \approx \frac{1}{2}m\bar{v}^2, \quad (20)$$

the relation (19) can also be written as

$$\bar{p} \approx \frac{2}{3}n\overline{\epsilon^{(k)}}. \quad (21)$$

- Note that (19) and (21) depend only on the number of molecules per unit volume, but make no reference to the nature of these molecules.
- These relations are thus equally valid, irrespective of whether the gas consists of molecules of He, Ne, O₂, N₂, or CH₄. The average pressure of any ideal gas kept in a container of fixed volume thus provides a very direct measure of the average kinetic energy of a molecule of the gas.

Numerical estimates

Before making some numerical estimates, it will be useful to recall some important definitions. The mass m of an atom or molecule is conveniently expressed in terms of some standard mass unit m_0 . In accordance with present international convention (adopted in 1960 and called the *unified scale of atomic weights*) this mass unit m_0 is defined in terms of the mass m_C of an atom of the particular carbon isotope ¹²C by‡

$$m_0 \equiv \frac{m_C}{12}. \quad (22)$$

† We neglect here the distinction between $\bar{v}^2$, the average of a square, and $\bar{v}^2$, the square of an average.

‡ We recall that a particular isotope is defined as an atom X with a uniquely specified nucleus; the symbol ⁿX indicates that the atom has n nucleons (protons + neutrons) in its nucleus. Atoms having nuclei with different numbers of neutrons but the same number of protons are chemically alike since they have the same number of extranuclear electrons.

The mass of a ^{12}C atom is thus exactly equal to 12 mass units. The mass of a H atom is then *approximately* equal to one mass unit.

The ratio of the mass m of an atom (or molecule) to the mass unit m_0 is called the *atomic weight* of the atom (or *molecular weight* of the molecule) and will be denoted by μ . Thus

$$\mu \equiv \frac{m}{m_0} \Rightarrow m = \mu \times m_0 \quad (23)$$

The atomic weight of ^{12}C is thus, by definition, equal to 12.

A convenient macroscopic number of atoms (or molecules) is the number N_a of atoms of mass m_0 which would have a total mass of 1 gram: i.e., the number N_a is defined by

$$N_a \equiv \frac{1}{m_0} \quad (24)$$

Alternatively, this defining relation can be written in the form

$$\begin{array}{l} 1(\text{atom}) \rightarrow m_0(\text{gm}) \\ N_a \rightarrow \mu(\text{gm}) \end{array} \quad N_a = \frac{m}{\mu m_0} = \frac{\mu}{m} \quad (25)$$

where we have used (23). This means that N_a is also equal to the number of molecules, of molecular weight μ , which have a total mass of μ grams. The number N_a is called *Avogadro's number*.

One *mole* of a certain kind of molecule (or atom) is defined as a quantity consisting of N_a molecules (or atoms) of this kind. A mole of molecules of molecular weight μ has, therefore, a total mass of μ grams.

The numerical value of Avogadro's number is experimentally found to be

$$N_a = (6.02252 \pm 0.00009) \times 10^{23} \text{ molecules/mole.} \quad (26)$$

(See the table of numerical constants at the end of the book.)

Let us now use the relations (19) or (21) for the pressure of a gas to estimate molecular quantities for nitrogen (N_2) gas, the main constituent of air. At room temperature and atmospheric pressure (10^6 dynes/cm²), the mass of N_2 gas contained within a vessel having a volume of one liter (10^3 cm³) is experimentally found to be about 1.15 grams. Since the atomic weight of a N atom is about 14, the molecular weight of a N_2 molecule is $2 \times 14 = 28$. It follows that 28 gm of N_2 gas contains Avogadro's number $N_a = 6.02 \times 10^{23}$ of N_2 molecules. The

total number N of molecules in the experimental vessel is thus

$$\begin{aligned} N &= (6.02 \times 10^{23}) \frac{1.15}{28} \\ &= 2.47 \times 10^{22} \text{ molecules} \end{aligned}$$

so that

$$\begin{aligned} n &\equiv \frac{N}{V} = \frac{2.47 \times 10^{22}}{10^3} \\ &\approx 2.5 \times 10^{19} \text{ molecules/cm}^3. \quad (27) \end{aligned}$$

Using (21), it then follows that the average kinetic energy of a N_2 molecule is

The equivalent mass of
(μ) an arbitrary atom
0 1 2 3 m_0
on the new scale.

$N_a \times m_0(\text{gm}) = 1(\text{gm})$
 $m_0 = \text{mass of one atom}$
• actually the mass of
one atom = $m(\text{gm})$
 $\Rightarrow N_a \times m(\text{gm}) = 1(\text{gm})$
 $1(\text{mole}) = \mu(\text{gm})$

$$\bar{\epsilon}^{(h)} \approx \frac{3}{2} \frac{\bar{p}}{n} = \frac{3}{2} \left(\frac{10^6}{2.5 \times 10^{19}} \right) \approx 6.0 \times 10^{-14} \text{ erg.} \quad (28)$$

$$m = \frac{28}{6.02 \times 10^{23}} = 4.65 \times 10^{-23} \text{ gm.} \quad (29)$$

Since N_0 nitrogen molecules (where N_0 is Avogadro's number) have a mass of 28 gm, the mass m of a single N_2 molecule is

$$\bar{v}^2 \approx \frac{2\bar{\epsilon}^{(h)}}{m} \approx \frac{2(6.0 \times 10^{-14})}{4.65 \times 10^{-23}} = 2.6 \times 10^9$$

$$\bar{v} \approx 5.1 \times 10^4 \text{ cm/sec.} \quad (30)$$

Mean free path

Focusing attention on a molecule in a gas at any instant of time, let us estimate the average distance l which this molecule travels before it collides with some other molecule in the gas. This distance l is called the *mean free path* of the molecule. To simplify matters, we can imagine that each molecule is spherical in shape and that the forces between any two molecules are similar to those between two hard spheres of radius a . This means that the molecules exert no force on each other as long as the separation R between their centers is greater than $2a$, but that they exert extremely large forces on each other (i.e., they *collide*) if $R < 2a$. Figure 1.35 illustrates an encounter between two such molecules. Here molecule A' can be considered stationary as molecule A approaches it with some relative velocity V in such a way that the centers of the molecules would approach within a distance b of each other if they remained undeflected. It is then apparent that the molecules will never collide if $b > 2a$, but that they will collide if $b < 2a$. Another way of expressing this geometrical relationship is to imagine that molecule A carries with it a disk of radius $2a$ (this disk being centered about the center of the molecule and oriented perpendicular to the velocity V). A collision between the two molecules will then occur only if the center of molecule A' lies within the volume swept out by the disk carried by molecule A .

The area σ of the imaginary disk carried by a molecule is

$$\sigma = \pi(2a)^2 = 4\pi a^2 \quad (31)$$

and is called the *total scattering cross section* for molecule-molecule collisions. The volume swept out by this disk as the molecule travels a distance l is equal to σl . Suppose that this volume is such that it contains, on the average, one other molecule, i.e., that

$$\langle \sigma l \rangle n \approx 1,$$

where n is the number of molecules per unit volume. Then the distance l is the average distance traveled by the molecule before suffer-

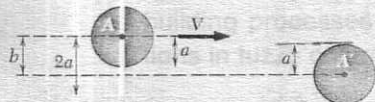


Fig. 1.35 Diagram illustrating an encounter between two hard spheres of radius a .

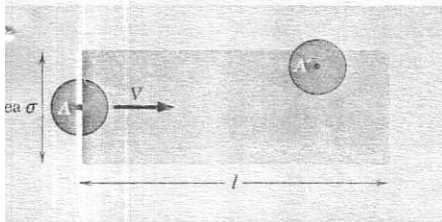


Fig. 1.36 Diagram illustrating the collisions suffered by a particular molecule A when it encounters another molecule A' whose center is located within the volume swept out by the area σ of the imaginary disk carried by A .

ing a collision with another molecule, i.e., it is the desired mean free path. Thus we get

$$l \approx \frac{1}{n\sigma}. \quad (32)$$

As one would expect, the mean free path becomes long (i) if n is small, so that there are only few molecules with which a given molecule can collide, and (ii) if the molecular radius is small, so that molecules must approach each other quite closely before they collide.

To estimate orders of magnitude, we return to the previously discussed case of N_2 gas at room temperature and atmospheric pressure. The radius of a molecule is of the order of 10^{-8} cm, i.e.,

$$a \sim 10^{-8} \text{ cm.}$$

Hence (31) gives for the cross section

$$\sigma = 4\pi a^2 \sim 12 \times 10^{-16} \text{ cm}^2.$$

Using the value (27) of n , (32) then yields the estimate

$$l \approx \frac{1}{n\sigma} \sim \frac{1}{(2.5 \times 10^{19})(12 \times 10^{-16})}$$

or $l \sim 3 \times 10^{-5} \text{ cm.} \quad (33)$

Note that $l \gg a$, i.e., the mean free path is much larger than the radius of a molecule. The molecules thus interact with each other infrequently enough so that it is a good approximation to treat the gas as ideal. On the other hand, the mean free path is very small compared to the linear dimensions of the one-liter vessel containing the gas.

1.7 Important Problems of Macroscopic Physics

Although the discussion of this chapter has been mostly qualitative, it has revealed the most significant features of macroscopic systems. We have, therefore, gained sufficient perspective to recognize some of the questions which we should ultimately like to explore and understand.

† Fundamental concepts

1. Our first task must clearly be that of transforming our qualitative insights into precisely formulated theoretical concepts capable of making quantitative predictions. For example, we have recognized that certain situations of a macroscopic system are more probable (or more random) than others. But precisely how does one assign a probability to a given macrostate of a system and how does one measure its degree of randomness? This question is of all-pervasive importance. We also concluded that the time-independent equilibrium situation is characterized as being the most random situation of an isolated system. The problem of defining randomness in a precise general way thus occurs again. Indeed, this problem proved troublesome to us when we tried to discuss the case of two arbitrary systems in thermal contact with

- each other. We suspected that the equilibrium condition of maximum randomness should imply that some parameter T (measuring roughly the mean energy per atom in a system) ought to be the same for both systems. But since we did not know how to define randomness in the general case, we were unable to arrive at an unambiguous definition of this parameter T (which we called the “absolute temperature”).
- (2) Thus we can pose the following basic question: How can we use probability ideas to describe macroscopic systems in a systematic way so as to define concepts such as randomness or absolute temperature?
- In discussing the example of the pendulum in Sec. 1.3, we saw that it is not readily apparent how to transform energy randomly distributed over many molecules to a less random form where it can do work by exerting a net macroscopic force over a macroscopic distance. This example illustrates questions of the most profound importance. Indeed, to what extent is it possible to take energy distributed randomly over many molecules of a substance (such as coal or gasoline) and to transform it into a less random form where it can be used to move pistons against opposing forces? In other words, to what extent is it possible to build the steam engines or gasoline engines responsible for our industrial revolution? Similarly, to what extent is it possible to take the energy distributed randomly over many molecules of certain chemical compounds and to transform it into a less random form where it can be used to produce muscular contraction or the synthesis of highly ordered polymer molecules such as proteins? In other words, to what extent can chemical energy be used to make biological processes possible? A good understanding of the concept of randomness should allow us to make significant statements about all these questions.

Properties of systems in equilibrium

Since macroscopic systems in equilibrium are particularly simple, their properties ought to be most readily amenable to quantitative discussion. There are, indeed, many equilibrium situations which are of great interest and importance. Let us mention some of the questions which are worthy of investigation.

- (1) • A homogeneous substance is one of the simplest systems whose equilibrium properties one may hope to calculate. For example, suppose that some particular fluid (gas or liquid) is in equilibrium at a given temperature. How does the magnitude of the pressure which it exerts depend on its temperature and on its volume? Or suppose that some substance contains a certain concentration of iron atoms, each of which has a definite magnetic moment. If this substance is at a given temperature and is located in a given magnetic field, what

$$P = P(T, V)$$

is the value of its magnetization, i.e., of its net magnetic moment per unit volume? How does this magnetization depend on the temperature and on the magnetic field? Or suppose that a small amount of heat is added to some particular substance (which might be a liquid, a solid, or a gas). By what amount does its temperature increase?

- Not only can we ask questions about macroscopic parameters of a system in equilibrium; we can also inquire into the behavior of the atoms of which it consists. For example, consider a container of gas maintained at some given temperature. All the molecules of the gas do not have the same speed and we can ask what fraction of the molecules have a speed in any specified range. If we make a very small hole in the container, some of the molecules will escape through the hole into the surrounding vacuum where their speed can be directly measured and the theory compared with the experimental results.
- Or consider an empty container whose walls are maintained at some elevated temperature. Since the atoms in the walls emit electromagnetic radiation, the container itself is then filled with radiation (or with photons) in equilibrium with the walls. What amount of energy of this electromagnetic radiation is concentrated in any given frequency range? If one makes a very small hole in the container through which some of the radiation escapes, then one can readily use a spectrometer to measure the amount of radiation emerging in any narrow frequency band and thus can compare the predictions with experimental results.
- This last problem is actually very important for understanding the radiation emitted by any hot body, whether it be the sun or the filament of a light bulb.
- Another type of situation of great interest is that where chemical reactions can occur between various kinds of molecules. As a specific example, consider a container of volume V which is filled with carbon dioxide (CO_2) gas. It is possible to transform CO_2 molecules into carbon monoxide (CO) and oxygen (O_2) molecules, or vice versa, according to the chemical reaction



- As the temperature of the container is raised, some of the CO_2 molecules will dissociate into CO and O_2 molecules. The container will then enclose a mixture of CO_2 , CO , and O_2 gases in equilibrium with each other. One would like to know how to calculate, from first principles, the relative number of CO_2 , CO , and O_2 molecules thus present in equilibrium at any specified temperature.

But even a simple substance consisting only of a single kind of molecule presents some intriguing problems. Such a substance can typi-

$$M = M(B, T)$$

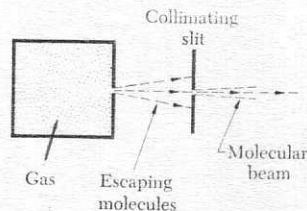


Fig. 1.37 A container filled with gas has a very small hole through which some molecules can escape into a surrounding vacuum. Collimation of these emerging molecules by one or more slits yields a well-defined molecular beam. The distribution of velocities of the molecules in the beam is closely related to that of the molecules in the container. Molecular beams of this sort represent a very powerful method for studying nearly isolated atoms or molecules and have been used to perform some of the most fundamental experiments of modern physics.

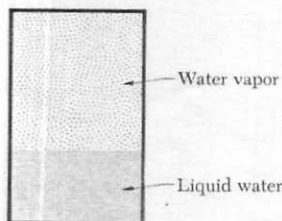


Fig. 1.38 Liquid water and its gas form, water vapor, are here shown coexisting in equilibrium at some specified temperature. The pressure exerted by the vapor then has a unique value which depends only on the temperature.

ally exist in several different forms, or *phases*, such as gas, liquid, or solid. A specific example might be water which can exist in the form of water vapor, liquid water, or ice. Each such phase consists of the same kind of molecule (H_2O , in the case of water), but the molecules are arranged differently. In the gas phase, the molecules are far apart from each other and thus move about almost independently of each other in random fashion. In the solid phase, on the other hand, the molecules are arranged in a very orderly manner. They are located in a regular crystal lattice near specific sites and are free merely to perform small oscillations about these sites. In the liquid phase the situation is intermediate, being neither as orderly as in the solid nor as random as in the gas. The molecules here are close to each other and always under each other's strong influence, yet they are free to move past each other over long distances. The evidence for these molecular arrangements in the various phases comes largely from studies of x-ray scattering.

It is well known that a substance changes from one phase to another phase at a very well-defined temperature (absorbing or giving off heat in the process). For example, water changes from the solid ice form to the liquid water form at 0° centigrade and changes in turn (at a pressure of 1 atmosphere) from the liquid water form to the gaseous steam form at 100° centigrade. Under certain circumstances two phases can, therefore, coexist in equilibrium with each other (e.g., ice and liquid water at 0° centigrade). A good theory of systems in equilibrium should allow us to make statements about the conditions of temperature and pressure under which two phases are thus able to coexist in equilibrium. It should also allow us to predict at what temperature a particular solid substance melts so as to become a liquid, and at what temperature the liquid vaporizes to become a gas.

These are actually very difficult and fascinating problems. The concept of degree of randomness, or order, is again crucial. As the absolute temperature (or mean energy per atom) of a substance is increased, it first changes from the most orderly (or least random) solid form to the liquid form which is of an intermediate degree of order; as the temperature is increased further, it changes from this liquid form to the gas form which is the most nearly random or disordered. But it is very striking that these changes from one degree of order to the next take place very abruptly at extremely well-defined temperatures. The essential reason is a kind of critical instability involving all the molecules of the substance. For example, suppose that the absolute temperature of a solid is sufficiently high so that its molecules, as a result of their relatively large mean energy, can oscillate about their

regular lattice positions with displacements large enough to be comparable with their intermolecular separation. Suppose then that there occurs a fluctuation in which a few adjacent molecules move simultaneously out of their regular lattice positions; this makes it easier for neighboring molecules slightly farther away also to move out of their regular positions; and so forth. The net result is then somewhat analogous to the collapse of a pile of dominoes; i.e., the high degree of order of the solid suddenly begins to disintegrate all at once and the solid turns into a liquid. This instability, which results in the melting of the solid, involves jointly all the molecules of the solid; it is, therefore, said to be a *cooperative phenomenon*. Since it does require the analysis of all the molecules in simultaneous interaction, the theoretical problem of treating any cooperative phenomenon such as melting or vaporization from a detailed microscopic point of view is very difficult and challenging.

Systems which are not in equilibrium

The discussion of systems which are not in equilibrium is ordinarily much more difficult than that of systems which are in equilibrium. Here it is necessary to deal with processes which change in time and the interest is focused on how fast or slowly they change. The discussion of such questions requires a detailed analysis of how effectively molecules interact with each other. Except in relatively simple cases such as dilute gases, this kind of analysis can become fairly complicated.

Let us mention again a few typical problems of interest. Consider, for example, the chemical reaction (34). Suppose that some CO_2 gas is introduced into the container at some given elevated temperature. How long would it take to achieve the equilibrium concentration of CO ? Thus one would like to calculate the rate at which the chemical reaction (34) proceeds from left to right.

As another illustration, consider two large bodies which are connected by a rod and which are at different temperatures T_1 and T_2 . Since this is not an equilibrium situation, heat will flow from one body to the other through the connecting rod. But how effectively is the energy transported down the rod, i.e., how long does it take for a given amount of heat to flow from one body to the other? This depends on an intrinsic property of the rod, its "thermal conductivity." For example, a rod of copper allows heat flow to proceed much more readily than a rod made of stainless steel; i.e., copper has a much higher thermal conductivity than stainless steel. The task of the theory is to define "thermal conductivity" precisely and to calculate this parameter from first principles.

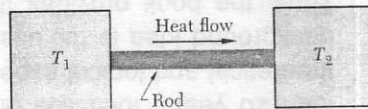


Fig. 1.39 Two bodies at different temperatures connected by a rod which conducts heat from one to the other.

Concluding Remarks

The preceding survey of problems indicates to some extent the range of macroscopic natural phenomena which we might like to treat quantitatively on the basis of fundamental microscopic considerations. We shall not learn how to discuss all such problems in the remainder of this book. Indeed, some of the questions which we have raised (e.g., calculations of phase transformations such as melting or vaporization), give rise to problems which are still not adequately solved and which are thus active areas of research. On the other hand, we are now well prepared to turn the qualitative considerations of this chapter into a more systematic quantitative discussion of the properties of macroscopic systems. This discussion will go a long way toward answering the most basic questions which we have raised.

Summary of Definitions

isolated system A system which does not interact with any other system.

ideal gas A gas of molecules whose mutual interaction is almost negligible (i.e., the interaction is large enough to ensure that the molecules can exchange energy among themselves, but is otherwise negligible).

ideal spin system A system of spins whose mutual interaction is almost negligible (i.e., the interaction is large enough to ensure that the spins can exchange energy among themselves, but is otherwise negligible).

microscopic Small, of the order of atomic dimensions or less.

macroscopic Very large compared to atomic dimensions.

microscopic state (or microstate) State of a system described in microscopic detail by the most complete specification, according to the laws of mechanics, of all the atoms in the system.

macroscopic state (or macrostate) State of a system described, without attention to microscopic details, by specifying only quantities which can be determined by macroscopic measurements.

macroscopic parameter A parameter which can be determined by large-scale measurements and which describes the macroscopic state of a system.

equilibrium A macroscopic state which does not tend to change in time, except for random fluctuations.

relaxation time Approximate time required for a system, starting from a situation far removed from equilibrium, to attain equilibrium.

irreversible process A process such that the time-reversed process (which would be observed in a movie played backward) would almost never occur in actuality.

thermal interaction Interaction which does not involve the performance of work on a macroscopic scale.

heat Energy transfer which is not associated with the performance of macroscopic work, but occurs on the atomic scale.

thermometer A small macroscopic system arranged so that only one of its macroscopic parameters changes when the system absorbs or gives off heat.

thermometric parameter The one variable macroscopic parameter of a thermometer.

temperature of a system with respect to a given thermometer The value of the thermometric parameter of the thermometer when the latter is placed in thermal contact with the system and allowed to come to equilibrium with it.

mean free path The average distance which a molecule travels in a gas before it collides with another molecule.

Suggestions for Supplementary Reading

F. J. Dyson, "What is Heat?", *Sci. American* 191, 58 (Sept. 1954).

R. Furth, "The Limits of Measurement," *Sci. American* 183, 48, (July 1950). A discussion of Brownian motion and other fluctuation phenomena.

B. J. Alder and T. E. Wainwright, "Molecular Motions," *Sci. American* 201, 113 (Oct. 1959).

This article discusses the application of modern high-speed computers to the study of molecular motions in various macroscopic systems.

Problems

1.1 Fluctuations in a spin system

Consider an ideal system of 5 spins in the absence of an external magnetic field. Suppose that one took a movie of this spin system in equilibrium. What fraction of the movie frames would show n spins pointing up? Consider all the possibilities $n = 0, 1, 2, 3, 4$, and 5.

1.2 Diffusion of a liquid

Suppose that a drop of a dye (having the same density as water) is introduced into a glass of water. The whole system is kept at a constant temperature and is left mechanically undisturbed. Suppose that one took a movie of the process occurring after the drop of dye had been put into the water. What would one see on a screen on which the movie is projected? What would one see if the movie is run backward through the projector? Is the process reversible or irreversible? Describe the process in terms of the motion of the dye molecules.

1.3 Microscopic explanation of friction

A wooden block, which has originally been given a push, is sliding on the floor and gradually comes to rest. Is this process reversible or irreversible? Describe the process as it would appear on a movie played backward. Discuss what happens during this process on the microscopic scale of atoms and molecules.

1.4 The approach to thermal equilibrium

Consider two gases A and A' in separate containers. Initially the average energy of a molecule of gas A is quite different from the average energy of a molecule of gas A'. The two containers are then placed in contact with each other so that energy in the form of heat can be transferred from gas A to the molecules of the container walls and thence to gas A'. Is the ensuing process reversible or irreversible? Describe in microscopic detail the process that would appear if one filmed the situation and ran this film backward through a projector.

1.5 Variation of gas pressure with volume

A container is divided into two parts by a partition. One of these parts has a volume V_i and is filled with a dilute gas; the other part is empty. Remove the partition and wait until the final equilibrium condition is attained where the molecules of the gas are uniformly distributed throughout the entire container of volume V_f .

(a) Has the total energy of the gas been changed? Use this result to compare the average energy per molecule and the average speed of a molecule in the equilibrium situations before and after the removal of the partition.

(b) What is the ratio of the pressure exerted by the gas in the final situation to that of the pressure exerted by it in the initial situation?

1.6 Number of gas molecules incident on an area

Consider nitrogen (N_2) gas at room temperature and atmospheric pressure. Using the numerical values given in the text, find the average number of N_2 molecules striking a 1-cm^2 area of the container walls per second.

1.7 Leak rate

A 1-liter glass bulb contains N_2 gas at room temperature and atmospheric pressure. The glass bulb, which is to be used in conjunction with some other experiment, is itself enclosed in a large evacuated chamber. Unfortunately the glass bulb has, unbeknown to the experimenter, a small pinhole about 10^{-5} cm in radius. To assess the importance of this hole, estimate the time required for 1 percent of the N_2 molecules to escape from the bulb into the surrounding vacuum.

1.8 Average time between molecular collisions

Consider nitrogen gas at room temperature and atmospheric pressure. Using the numerical values given in the text, find the average time a N_2 molecule travels before colliding with another molecule.

***1.9 Equilibrium between atoms of different masses**

Consider a collision between two different atoms having masses m_1 and m_2 . Denote the velocities of these atoms before the collision by $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively; denote their velocities after the collision by $\mathbf{v}_1'$ and $\mathbf{v}_2'$, respectively. It is of interest to investigate the energy transferred from one atom to the other as a result of the collision.

(a) Introduce the relative velocity $\mathbf{V} \equiv \mathbf{v}_1 - \mathbf{v}_2$ and the center-of-mass velocity $\mathbf{c} \equiv (m_1\mathbf{v}_1 + m_2\mathbf{v}_2)/(m_1 + m_2)$. The relative velocity after the collision is then $\mathbf{V}' = \mathbf{v}_1' - \mathbf{v}_2'$. In a collision $\mathbf{c}$ remains unchanged by virtue of conservation of momentum, while $|\mathbf{V}'| = |\mathbf{V}|$ by virtue of conservation of energy. Show that the energy gain $\Delta\epsilon_1$ of atom 1 in a collision is given by

$$\Delta\epsilon_1 = \frac{1}{2}m_1(v_1'^2 - v_1^2) = m_1m_2(m_1 + m_2)^{-1}\mathbf{c} \cdot (\mathbf{V}' - \mathbf{V}). \quad (\text{i})$$

(b) Denote by θ the angle between $\mathbf{V}'$ and $\mathbf{V}$, by φ the angle between the plane containing $\mathbf{V}'$ and $\mathbf{V}$ and the plane containing $\mathbf{c}$ and $\mathbf{V}$, and by ψ the angle between $\mathbf{c}$ and $\mathbf{V}$. Show that (i) then becomes

$$\Delta\epsilon_1 = m_1m_2(m_1 + m_2)^{-1}cV[(\cos\theta - 1)\cos\psi + \sin\theta\sin\psi\cos\varphi] \quad (\text{ii})$$

where $cV\cos\psi = \mathbf{c} \cdot \mathbf{V} = (m_1 + m_2)^{-1}[m_1v_1^2 + (m_2 - m_1)\mathbf{v}_1 \cdot \mathbf{v}_2 - m_2v_2^2]$.

(c) Consider two atoms of this kind in a gas where many collisions occur. On the average, the azimuthal angle φ is then as often positive as negative so that $\overline{\cos\varphi} = 0$; also, since $\mathbf{v}_1$ and $\mathbf{v}_2$ have random directions, the cosine of the angle between them is as often positive as negative so that $\overline{\mathbf{v}_1 \cdot \mathbf{v}_2} = 0$. Show that (ii) therefore becomes, on the average,

$$\overline{\Delta\epsilon_1} = \frac{2m_1m_2}{(m_1 + m_2)^2} (\overline{1 - \cos\theta})(\bar{\epsilon}_2 - \bar{\epsilon}_1) \quad (\text{iii})$$

where $\epsilon_1 = \frac{1}{2}m_1v_1^2$ and $\epsilon_2 = \frac{1}{2}m_2v_2^2$.

In the equilibrium situation in particular the energy of an atom must, on the average, remain unchanged so that $\overline{\Delta\epsilon_1} = 0$. Show that (iii) then implies that

$$\text{in equilibrium,} \quad \bar{\epsilon}_2 = \bar{\epsilon}_1. \quad (\text{iv})$$

Thus the average energies of interacting atoms in equilibrium are equal even if the masses of the atoms are different.

1.10 Comparison of molecular speeds in a gas mixture

Consider a gas mixture enclosed in some container and consisting of monatomic molecules of two different masses m_1 and m_2 .

(a) Suppose that this gas mixture is in equilibrium. Use the result of the preceding problem to find the approximate ratio of the average speed $\bar{v}_1$ of a molecule of mass m_1 to the average speed $\bar{v}_2$ of a molecule of mass m_2 .

(b) Suppose that the two kinds of molecules are He (helium) and Ar (argon) which have atomic weights equal to 4 and 40, respectively. What is the ratio of the average speed of a He atom to that of an Ar atom?

1.11 Pressure of a gas mixture

Consider an ideal gas which consists of two kinds of atoms. To be specific, suppose that there are, per unit volume, n_1 atoms of mass m_1 and n_2 atoms of mass m_2 . The gas is assumed to be in equilibrium so that the average energy $\bar{\epsilon}$ per atom is the same for both kinds of atoms. Find an approximate relation for the average pressure $\bar{p}$ exerted by the gas mixture. Express your result in terms of $\bar{\epsilon}$.

1.12 Mixing of two gases

Consider a container divided into two equal parts by a partition. One of these parts contains 1 mole of helium (He) gas, the other 1 mole of argon (Ar) gas. Energy in the form of heat can pass through the partition from one gas to the other. After a sufficiently long time, the two gases will, therefore, come to equilibrium with each other. The average pressure of the helium gas is then $\bar{p}_1$, that of the argon gas is $\bar{p}_2$.

(a) Compare the pressures $\bar{p}_1$ and $\bar{p}_2$ of the two gases.

(b) What happens when the partition is removed? Describe the process as it would appear on a movie played backward. Is it reversible or irreversible?

(c) What is the average pressure exerted by the gas in the final equilibrium situation?

1.13 Effect of a semipermeable partition ("osmosis")

A glass bulb contains argon (Ar) gas at room temperature and at a pressure of 1 atmosphere. It is placed in a large chamber containing helium (He) gas also at room temperature and at a pressure of 1 atmosphere. The bulb is made of a glass which is permeable to the small He atoms, but is impermeable to the larger Ar atoms.

(a) Describe the process which ensues.

(b) What is the most random distribution of the molecules which is attained in the ultimate equilibrium situation?

(c) What is the average pressure of the gas inside the bulb when this ultimate equilibrium situation has been attained?

1.14 Thermal vibrations of the atoms in a solid

Consider nitrogen (N_2) gas in equilibrium within a box at room temperature. In accordance with the result of Prob. 1.9, it is then reasonable to assume that the average kinetic energy of a gas molecule is roughly equal to the average kinetic energy of an atom in the solid wall of the container. Each atom in such a solid is localized near a fixed site. It is, however, free to oscillate about this site and should, to good approximation, perform simple harmonic motion about this position. Its potential energy is then, on the average, equal to its kinetic energy.

Suppose that the walls consist of copper which has a density of 8.9 gm/cm^3 and an atomic weight of 63.5.

(a) Estimate the average speed with which a copper atom vibrates about its equilibrium position.

(b) Estimate roughly the mean spacing between the copper atoms in the solid. (You may assume them to be located at the corners of a regular cubic lattice.)

(c) When a force F is applied to a copper bar of cross-sectional area A and length L , the increase in length ΔL of the bar is given by the relation

$$\frac{F}{A} = Y \frac{\Delta L}{L}$$

where the proportionality constant Y is called *Young's modulus*. Its measured value for copper is $Y = 1.25 \times 10^{12} \text{ dynes/cm}^2$. Use this information to estimate the restoring force acting on a copper atom when it is displaced by some small amount x from its equilibrium position in the solid.

(d) What is the potential energy of an atom when it is displaced by an amount x from its equilibrium position? Use this result to estimate the average magnitude $|x|$ of the amplitude of vibration of a copper atom about its equilibrium position. Compare $|x|$ with the separation between the copper atoms in the solid.

Chapter 2

Basic Probability Concepts

- 2.1	Statistical Ensembles	56
2.2	Elementary Relations among Probabilities	64
- 2.3	The Binomial Distribution	67
2.4	Mean Values	75
- 2.5	Calculation of Mean Values for a Spin System	80
2.6	Continuous Probability Distributions	86
	Summary of Definitions	90
	Important Relations	90
	Suggestions for Supplementary Reading	91
	Problems	91

Probability is a distance function
of the event Space E , that is to say Probability
is the Metric d of the event Space $\Rightarrow$
If $x, y \in E \Rightarrow d(x, y)$ determines the
distance between the occurrence of either
 x or $y \equiv$ Probability.

Chapter 2 Basic Probability Concepts

The considerations of the preceding chapter have indicated that probability arguments are of fundamental significance for the understanding of macroscopic systems consisting of very many particles. It will, therefore, be useful to review the most basic notions of probability and to examine how they can be applied to some simple, but important, problems. Indeed, this discussion is likely to be valuable in contexts far wider than those of immediate interest to us. For example, probability concepts are indispensable in all games of chance, in the insurance business (with its need to assess the probable occurrence of death or disease among its clients), and in sampling procedures such as public opinion polls. In biology, they are profoundly important in the study of genetics. In physics, they are needed in treating the occurrence of radioactive decay, the arrival of cosmic rays at the surface of the earth, or the random emission of electrons from the hot filament of a vacuum tube; furthermore, they play a fundamental role in the quantum-mechanical description of atoms and molecules. Most important to us, they will form the basis of our entire discussion of macroscopic systems.

2.1 Statistical Ensembles

Consider a system A on which one can perform experiments or observations.[†] In many cases the particular outcome resulting from the performance of a single experiment cannot be predicted with certainty, either because this is intrinsically impossible[‡] or because the information available about the system is insufficient to permit such a unique prediction. Although it is not possible to make statements about a single experiment, it may still be possible to make significant statements about the outcomes of a large number of similar experiments. We are thus led to a *statistical* description of the system, i.e., to a

[†] An act of observation can be regarded as an experiment in which the result of the observation constitutes the outcome of the experiment. Therefore we need make no distinction between experiments and observations.

[‡] This is the situation, for example, in quantum mechanics where the outcome of a measurement on a microscopic system is ordinarily not predictable with certainty.

description in terms of *probabilities*. To achieve such a description, we proceed in the following way:

Instead of focusing attention on the single system A of interest, we contemplate an assembly (or an *ensemble*, in more customary terminology) consisting of some very large number $\mathcal{N}$ of "similar" systems. In principle, $\mathcal{N}$ is imagined to be arbitrarily large (i.e., $\mathcal{N} \rightarrow \infty$). The systems are supposed to be "similar" in the sense that each system satisfies the same conditions known to be satisfied by the system A . This means that each system is imagined to have been prepared in the same way as A and to be subjected to the same experiment as A . We can then ask in what fraction of cases a particular outcome of the experiment does occur. To be precise, we arrange matters so that we can enumerate in some convenient way all the possible mutually exclusive outcomes of the experiment. (The total number of such possible outcomes may be finite or infinite.) Suppose then that a particular outcome of the experiment is labeled by r and that there are, among the $\mathcal{N}$ systems of the ensemble, $\mathcal{N}_r$ systems which exhibit this outcome. Then the fraction

$$P_r \equiv \frac{\mathcal{N}_r}{\mathcal{N}} \quad (\text{where } \mathcal{N} \rightarrow \infty) \quad (1)$$

is called the *probability of occurrence of the outcome r* . To the extent that $\mathcal{N}$ is made very large, a repetition of the same experiment on the ensemble is expected to lead with increasing reproducibility to the same ratio $\mathcal{N}_r/\mathcal{N}$. The definition (1) becomes thus unambiguous in the limit where $\mathcal{N}$ is made arbitrarily large.

The preceding discussion shows how the probability of occurrence of any possible outcome of an experiment on a system can be measured by repeating the experiment on a large number $\mathcal{N}$ of similar systems.† Although the outcome of the experiment on a single system cannot be predicted, the task of a *statistical* theory is then that of predicting the *probability* of occurrence of each of the possible outcomes of the experiment. The predicted probabilities can then be compared against the probabilities actually measured by experiments on an ensemble of similar systems.

Several examples will serve to illustrate this statistical description in a number of concrete cases.

† If the situation being contemplated is independent of time, one could equally well repeat the same experiment $\mathcal{N}$ times in succession with the *one* particular system under consideration (being careful to start the experiment each time with the system in the same initial condition).

Can be represented as
a Convergent Sequence.

* Both Theory and
Experiment work with $\mathcal{N}$.

Statistical Theory
Predictions are Compared
against Experimental
Outcomes.

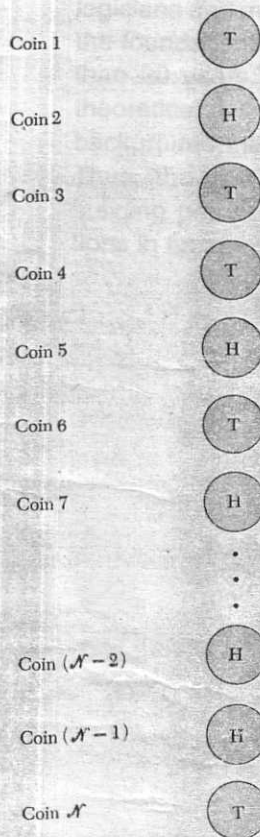


Fig. 2.1 In order to make probability statements about the tossing of a single coin, one considers an ensemble consisting of a large number N of similar coins. The diagram indicates schematically the appearance of this ensemble after each coin has been tossed. The symbol H denotes "head," the symbol T "tail."

Tossing of coins or dice

Consider the experiment of tossing a coin. There are only two possible outcomes of such an experiment, either "head" or "tail," depending on whether the coin comes to rest on a table with the engraved head uppermost or its other side uppermost.[†] In principle, the outcome of the experiment could be completely predicted if we knew exactly how the coin is tossed and the forces by which it interacts with the table. It would then only be necessary to make the requisite complicated calculations based on the laws of classical mechanics. But such detailed information about the tossing of the coin is not available in practice. It is, therefore, not possible to make a unique prediction about the outcome of a particular toss. (Indeed, even if all the requisite information could be obtained and all the requisite computations could be performed, we would ordinarily not be interested in making precise predictions at the expense of such enormous complexity.) The statistical formulation of the experiment is, however, very easy and familiar. We need only consider an ensemble consisting of a very large number N of similar coins. When these coins are tossed in similar ways, we can count the fraction of cases where the outcome is a head and that where it is a tail.[‡] These fractions then give, respectively, the measured probability p of obtaining a head and probability q of obtaining a tail. A statistical theory would attempt to predict these probabilities. For example, if the center of mass of the coin coincides with its geometrical center, the theory might be based on the

symmetry argument that there is nothing in the laws of mechanics which distinguishes between a head or a tail. In this case, half the experimental outcomes should be heads and half tails, so that $p = q = \frac{1}{2}$. Comparison with experiment might or might not verify the theory. For example, if heads were observed to occur more frequently than tails, we might conclude that the theory is not justified in assuming that the center of mass of the coin coincides with its geometrical center.

Consider now the somewhat more complicated experiment of tossing a set of N coins. Since the tossing of any one coin can have 2 possible outcomes, the tossing of the set of N coins can then have any of $2 \times 2 \times \cdots \times 2 = 2^N$ possible outcomes.[§] The statistical formulation of the experiment requires again that, instead of dealing with a single set of N coins, we contemplate an ensemble consisting of M such sets of N coins, where each set is tossed in a similar way. A possible question of interest is then the probability that any particular one of the 2^N possible outcomes occurs in the ensemble. A less detailed question of interest might concern the probability of finding in the ensemble an outcome where any n of the coins show heads and the remaining $(N - n)$ coins show tails.

The problem of tossing a set of N dice is, of course, analogous. The only difference is that the toss of any one die can have 6 possible outcomes, depending on which of the six faces of the cubical die lands uppermost.

It is often convenient to use the simple word *event* to denote the outcome of an experiment or the result of an observation. Note that the probability of occurrence of an event depends crucially on the information available about the system under consideration. Indeed,

[†] We neglect the remote possibility that the coin comes to rest standing on its edge.

[‡] Alternatively, we might take the same coin and toss it N times in succession, counting the fraction of cases showing heads or tails.

[§] In the special case where $N = 4$, the $2^4 = 16$ possible outcomes are listed explicitly in Table 1.1 (p. 7) if one interprets the letter L as a head and the letter R as a tail.

this information determines the kind of statistical ensemble to be contemplated, since this ensemble must consist solely of systems which satisfy all the conditions satisfied by the particular system under consideration.

Example

Suppose that we are interested in the following question: What is the probability that a person living in the United States will be hospitalized at some time between the ages of 23 and 24 years? Then we must consider an ensemble consisting of a large number of persons in the United States and must ascertain the fraction of these persons hospitalized at some time between the ages of 23 and 24. But suppose that we are now also

told that the person is of the female sex. Then the answer to our question changes because we must now contemplate an ensemble of *women* living in the United States and must ascertain how many of these women are hospitalized at some time between the ages of 23 and 24. (Indeed, women of this age tend to enter hospitals because of childbirth, a condition from which men are exempt.)

Application to systems of many particles

Consider a macroscopic system consisting of many particles. For example, the system might be an ideal gas of N molecules, a system of N spins, a liquid, or a piece of copper. In none of these cases is it possible to make a unique prediction about the behavior of each of the particles in the system,† nor is it of interest. We resort, therefore, to a statistical description of the system A under consideration. Instead of considering the single system A , we contemplate an ensemble consisting of a large number $\mathcal{N}$ of systems similar to A . To make a statistical statement about the system at the time t , we make observations on the $\mathcal{N}$ systems at this time t . Thus we can determine the probability $P_r(t)$ that the observation yields a particular outcome r at the time t . The procedure is most readily visualized by imagining that we take a motion picture of each system in the ensemble. We then end up with $\mathcal{N}$ film strips which contain the results of all observations on the systems of the ensemble. The behavior of any one system in the ensemble, say system number k , as a function of time can then be obtained by looking at the k th film strip (i.e., by looking along a particular horizontal line in Fig. 2.3). On the other hand,

† In the correct quantum-mechanical description of the system nonstatistical predictions are not possible, even in principle. In a classical description a unique prediction about a system would require knowledge of the position and velocity of each particle at one time, information which is not available to us.

$$\left\{ \left\{ (C_{ijk})_{i=1}^N \right\}_{j=1}^{2^N} \right\}_{k=1}^{\mathcal{N}} = \Omega$$

Here,

C_{ijk} "is a face of a coin"

H only or T only

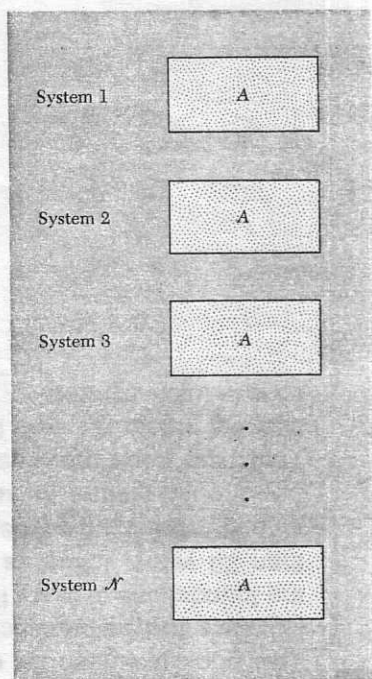


Fig. 2.2 Statistical description of a system A consisting of a gas in a box. The diagram illustrates schematically a statistical ensemble of $\mathcal{N}$ such systems similar to the system A under consideration.

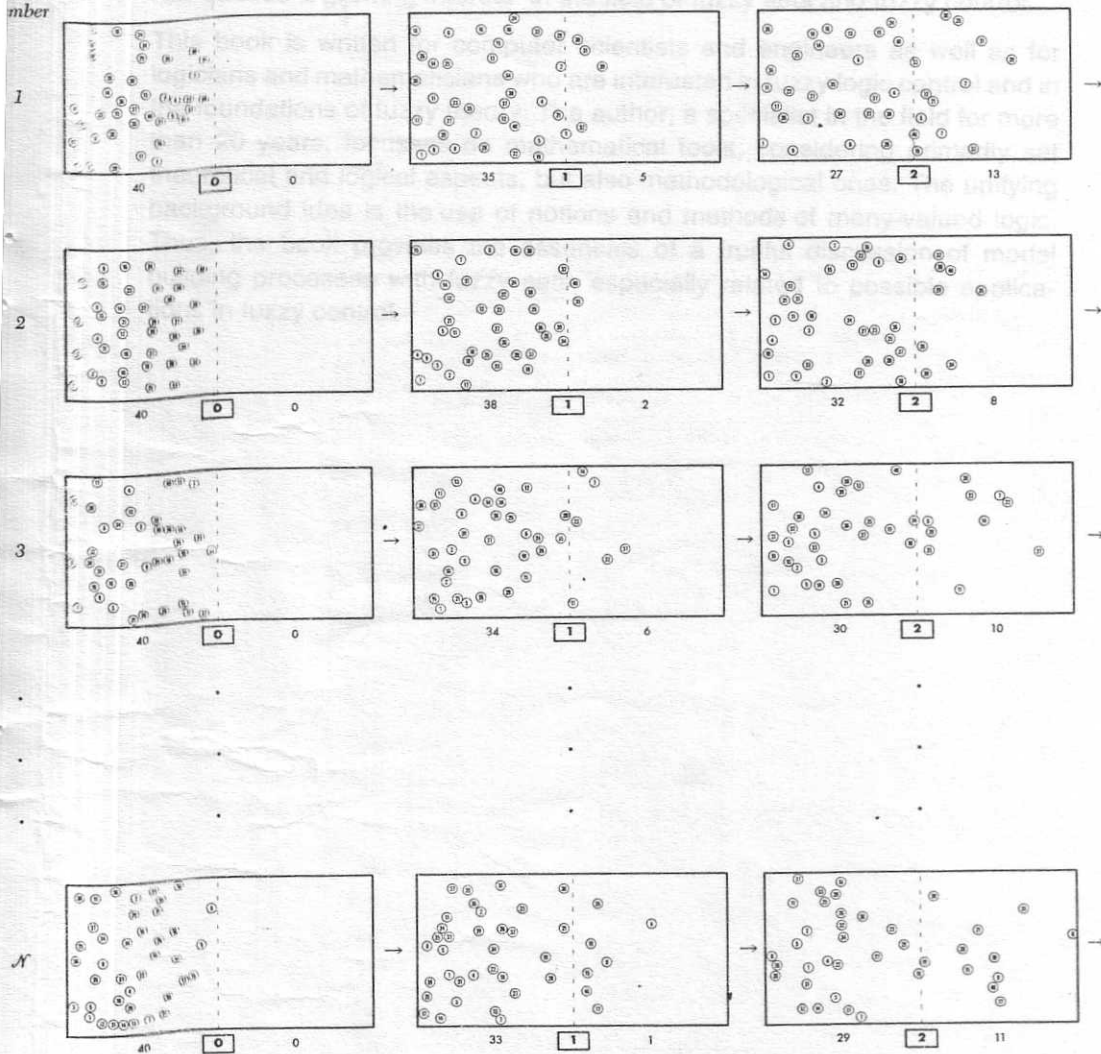


Fig. 2.3 Computer-made pictures showing a statistical ensemble of systems. This ensemble was constructed to represent a system consisting of 40 particles in a box when the information available about the system is the following: All particles are

known to be in the left half of the box at some initial time corresponding to the frame $j = 0$, but nothing else is known about their positions or velocities.

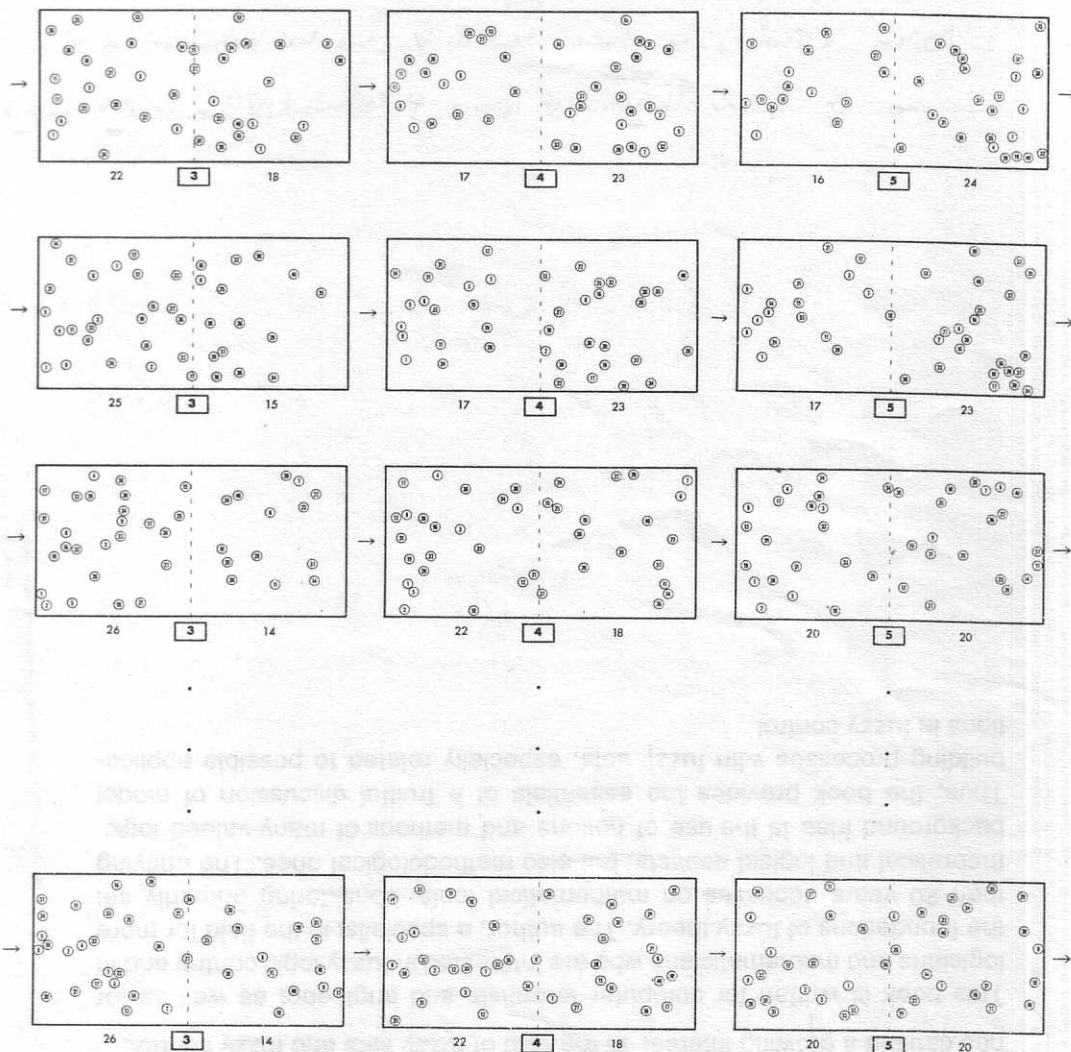


Fig. 2.3 (cont.)

The evolution in time of the k th system in the ensemble can be followed by looking horizontally at the successive frames $j = 0, 1, 2, \dots$ for this system. Statistical statements about the

system at any time corresponding to the j th frame can be made by looking vertically at all the systems in their j th frame and doing the counting necessary to determine probabilities.

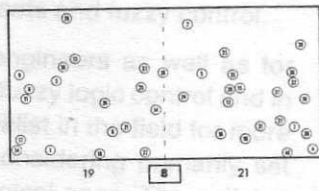
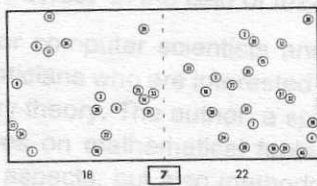
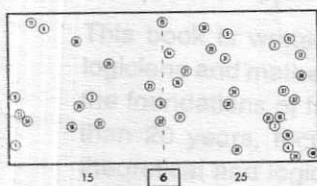
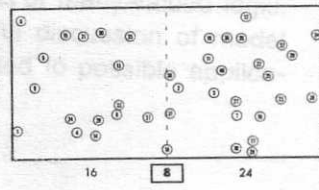
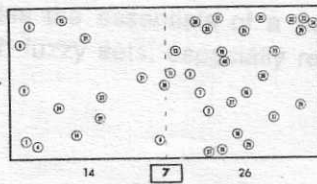
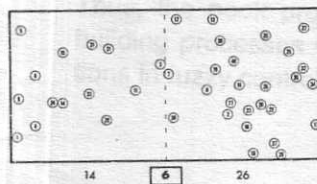
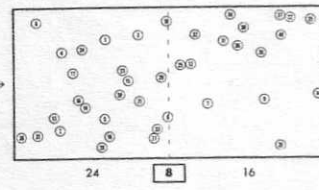
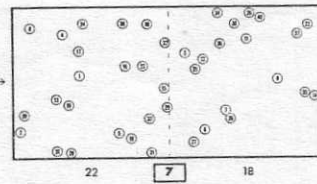
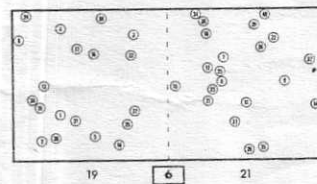
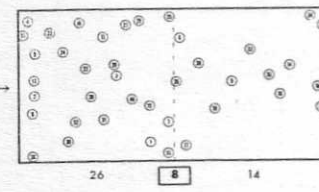
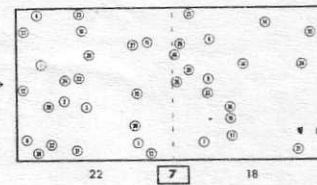
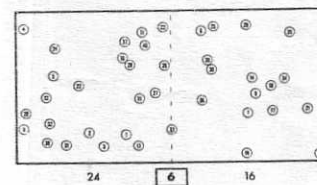
1
→2
→3
→N
→

Fig. 2.4 Continuation of Fig. 2.3. The ensemble has by now become time-independent, i.e., the system has attained equilibrium.

1 System (George) in some potential V and has behavior in V

$\rightarrow$ if $P_r \sim$ ^{State} (behavior) of George & equivalent V 's $\rightarrow$ Equilibrium state

probability statements about the system at any one time t are obtained by looking at all the film frames taken at the particular time t (i.e., by looking along a particular vertical line in Fig. 2.3 and by counting the fraction of systems exhibiting a given outcome at this time).

A statistical ensemble of systems is said to be time-independent if the number of its systems exhibiting any particular event is the same at every time (or equivalently, if the probability of occurrence of any particular event in this ensemble is independent of time). The statistical description then provides a very clear definition of equilibrium: *An isolated macroscopic system is said to be in equilibrium if a statistical ensemble of such systems is time-independent.*

Example

Consider an ideal gas of N molecules. At some initial time t_0 , immediately after a partition has been removed, all the molecules of this gas are known to be in the left half of a box. How do we go about giving a statistical description of what happens at all subsequent times? We need only consider an ensemble consisting of some large number $\mathcal{N}$ of similar boxes of gas, each one of which has all its molecules concentrated in its left half at the time t_0 . Such an ensemble is shown schematically in Fig. 2.3. We can then consider this ensemble at any time $t > t_0$ and ask various questions of interest. For example, focusing attention on any one molecule, what is the probability $p(t)$ that this molecule is in the left half of the container or the probability $q(t)$ that it is in the right half of the container? Or what is the probability $P(n, t)$ that, at any time t , n of the N molecules are located in the left half of the container? At the initial time t_0 , we know that $p(t_0) = 1$ while $q(t_0) = 0$.

[Similarly, $P(N, t_0) = 1$, while $P(n, t_0) = 0$ for $n \neq N$.] As time goes on, all these probabilities change until the molecules become uniformly distributed throughout the box so that $p = q = \frac{1}{2}$. Thereafter the probabilities remain unchanged in time, i.e., the ensemble has become time-independent and the system has attained equilibrium (see Fig. 2.4).† This time-independent situation is, of course, particularly simple. Indeed, the problem of the gas of N molecules then becomes analogous to the previously discussed problem of the set of N coins. In particular, the probability p of finding a molecule in the left half of the box is analogous to the probability p that a tossed coin shows a head; similarly, the probability q of finding a molecule in the right half of the box is analogous to the probability q that a coin shows a tail. Furthermore, as in the case of the coin, these probabilities are time-independent and such that $p = q = \frac{1}{2}$.

† Note that, despite the erratic fluctuations occurring in any single system as time goes on, the probability P in the ensemble at any one instant of time has always a unique well-defined value since the number $\mathcal{N}$ of systems in the ensemble is arbitrarily large. This remark illustrates the great conceptual simplicity of thinking in terms of ensembles rather than in terms of a single system.

For any r ,

$$P_r = \text{Constant}, \forall t$$

Remark

In Eq. 1.4 of Chap. 1 we calculated probabilities by considering only a single system. This is ordinarily a valid procedure in the special case of a system in equilibrium. Since an ensemble of such systems is time-independent, a large number of successive observations on a single system is then equivalent to a large number of simultaneous observations on the many systems of the ensemble. In other words, suppose that one

takes a film strip of a single system over a time τ and cuts it into $\mathcal{N}$ long pieces, each one of duration $\tau_1 \equiv \tau/\mathcal{N}$ (where τ_1 is so long that the behavior of the system in one piece is independent of its behavior in an adjacent piece). Then the collection of these $\mathcal{N}$ film strips of a single system is indistinguishable from a collection of $\mathcal{N}$ film strips taken of all the systems in the ensemble during any time interval of duration τ_1 .

2.2 Elementary Relations among Probabilities

Probabilities satisfy some simple relations which are almost self-evident, but quite important. It will be worth deriving these relations by starting directly from the definition (1) of a probability. Throughout the following discussion the number $\mathcal{N}$ of systems in the ensemble will always be understood to be infinitely large.

Suppose that experiments on some system A can lead to any of α mutually exclusive outcomes. Let us label each such outcome or *event* by some index r ; the index r can then refer to any of the α numbers $r = 1, 2, 3, \dots$, or α . In an ensemble of similar systems, $\mathcal{N}_1$ of them will be found to exhibit event 1, $\mathcal{N}_2$ of them event 2, $\dots$, and $\mathcal{N}_\alpha$ of them event α . Since these α events are mutually exclusive and exhaust all possibilities, it follows that

$$\mathcal{N}_1 + \mathcal{N}_2 + \dots + \mathcal{N}_\alpha = \mathcal{N}.$$

Dividing by $\mathcal{N}$, this relation becomes

$$P_1 + P_2 + \dots + P_\alpha = 1 \quad (2)$$

where $P_r \equiv \mathcal{N}_r/\mathcal{N}$ denotes the probability of occurrence of event r in accordance with the definition (1). The relation (2), which states merely that the sum of all possible probabilities adds up to unity, is called the *normalization condition* for probabilities. Using the summation symbol Σ defined in (M.1), this relation can also be written more compactly as

$$\sum_{r=1}^{\alpha} P_r = 1. \quad (3)$$

What is the probability of occurrence of *either* event r or event s ? There are $\mathcal{N}_r$ systems in the ensemble exhibiting event r and $\mathcal{N}_s$ systems exhibiting event s . Hence there are $(\mathcal{N}_r + \mathcal{N}_s)$ systems exhibit-

ing either event r or event s . Correspondingly the probability $P(r \text{ or } s)$ of occurrence of either alternative, event r or s , is simply given by

$$P(r \text{ or } s) = \frac{\mathcal{N}_r + \mathcal{N}_s}{\mathcal{N}}$$

so that

$$P(r \text{ or } s) = P_r + P_s. \quad (4)$$

Example

Suppose that we consider the tossing of a die which, by virtue of its symmetry, has equal probabilities $\frac{1}{6}$ of landing with any of its six faces uppermost. The probability that the die exhibits the number 1 is then $\frac{1}{6}$, as is the probability that it exhibits the number 2. The probability

that the die exhibits either the number 1 or the number 2 is then, by (4), simply $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. This is, of course, an obvious result since the events where 1 or 2 are uppermost represent one-third of all the six possible events 1, 2, 3, 4, 5, or 6.

The relation (4) is immediately generalizable to more than two alternative events. Thus the probability of occurrence of either one of several events is simply equal to the sum of their respective probabilities. In particular, we see that the normalization condition (2) merely states the obvious result that the sum of the probabilities on the left (i.e., the probabilities of occurrence of *either* event 1 *or* event 2 . . . *or* event α) is simply equal to unity (i.e., equivalent to certainty) since all possible events are exhausted in the enumeration of the α possible alternatives.

Joint probabilities

Suppose that the system under consideration can exhibit two different types of events, namely α possible events of a type labeled by r (where the index $r = 1, 2, 3, \dots, \alpha$) and β possible events of a type labeled by s (where the index $s = 1, 2, 3, \dots, \beta$). We shall denote by P_{rs} the probability of joint occurrence of *both* event r *and* event s . That is, in an ensemble consisting of a large number $\mathcal{N}$ of similar systems, $\mathcal{N}_{rs}$ of these are characterized by the joint occurrence of *both* an event r of the first type *and* an event s of the second type. Then $P_{rs} \equiv \mathcal{N}_{rs}/\mathcal{N}$. As usual, we denote by P_r the probability of occurrence of an event r (irrespective of the occurrence of events of type s). That is, if in the previous ensemble one pays no attention to events of type s and counts $\mathcal{N}_r$ systems in the ensemble exhibiting event r , then $P_r \equiv \mathcal{N}_r/\mathcal{N}$. Similarly, we denote by P_s the probability of occurrence of an event s (irrespective of the occurrence of events of type r).

Conditional Probability

$$P(R|Q) = P(R \cap Q) / P(Q)$$

$$P(Q|R) = P(Q \cap R) / P(R)$$

$$\Rightarrow P(R \cap Q) = P(R|Q) P(Q)$$

$$= P(Q|R) P(R)$$

For independent R and Q

either the occurrence of

R or Q does not give

information at all about

the occurrence of Q or R

$\Rightarrow P(Q|R)$ becomes $P(Q)$

$\& P(R|Q)$ becomes $P(R)$

and $P(R \cap Q) = P(R) P(Q)$

A special, but important, case is that where the probability of occurrence of an event of type s is unaffected by the occurrence or nonoccurrence of an event of type r . The events of type r and s are then said to be *statistically independent* or *uncorrelated*. Now consider in the ensemble the $\mathcal{N}_r$ systems which exhibit any particular event r . Irrespective of the particular value of r , a fraction P_s of these systems will then also exhibit the event s . Thus the number $\mathcal{N}_{rs}$ of systems exhibiting jointly both r and s is simply

$$\mathcal{N}_{rs} = \mathcal{N}_r P_s.$$

Correspondingly the probability of joint occurrence of both r and s is given by

$$P_{rs} = \frac{\mathcal{N}_{rs}}{\mathcal{N}} = \frac{\mathcal{N}_r P_s}{\mathcal{N}} = P_r P_s.$$

Hence we conclude that

if the events r and s are statistically independent,

$$P_{rs} = P_r P_s.$$

(5)

Note that the result (5) is not true if the events r and s are not statistically independent. The relation (5) can be immediately generalized; thus the joint probability of more than two statistically independent events is simply the product of their respective probabilities.

Example

Suppose that the system A under consideration consists of two dice A_1 and A_2 . An event of type r might be the appearance uppermost of any one of the 6 faces of the die A_1 ; similarly an event of type s might be the appearance uppermost of any one of the 6 faces of die A_2 . An event of system A is then specified by stating which face of die A_1 is uppermost and which face of die A_2 is uppermost. An experiment involving the tossing of the two dice would thus have $6 \times 6 = 36$ possible outcomes. Probability statements can be made about an ensemble consisting of a large number $\mathcal{N}$ of similar pairs of dice. We suppose that each die is symmetric so that it is equally likely to land with any one of its 6 faces upper-

most. The probability P_r that each die lands with any particular face r uppermost is then simply $\frac{1}{6}$. If the dice do not interact with each other (e.g., if they are not both magnetized so as to exert forces tending to align each other) and are not tossed in precisely identical ways, they can be considered statistically independent. In this case the joint probability P_{rs} that die A_1 lands with a particular face r uppermost and that die A_2 lands with a particular face s uppermost is simply

$$P_{rs} = P_r P_s = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}.$$

This result is, of course, quite evident since the contemplated event represents one out of $6 \times 6 = 36$ possible outcomes.

2.3 The Binomial Distribution

We have now acquired sufficient familiarity with statistical methods to give a quantitative discussion of some physically important problems. Consider, for instance, an ideal system of N spins $\frac{1}{2}$, each having an associated magnetic moment μ_0 . This system is of particular interest since it is a very simple system readily described in terms of quantum mechanics; it will therefore often be useful as a prototype of more complicated systems. For the sake of generality, we suppose that the spin system is located in an external magnetic field $\mathbf{B}$. Each magnetic moment can then point either "up" (i.e., parallel to field $\mathbf{B}$), or "down" (i.e., antiparallel to the field $\mathbf{B}$). We assume that the spin system is in equilibrium. A statistical ensemble consisting of $\mathcal{N}$ such spin systems is thus time-independent. Focusing attention on any one spin, we shall denote by p the probability that its magnetic moment points up and by q the probability that it points down. Since these two orientations exhaust all the possibilities, the normalization requirement (3) has the obvious consequence that

$$p + q = 1 \quad (6)$$

or $q = 1 - p$. In the absence of the field, when $\mathbf{B} = 0$, there is no preferred direction in space so that $p = q = \frac{1}{2}$. But in the presence of the field, a magnetic moment will be more likely to point along the field than opposite to it so that $p > q$.† Since the spin system is ideal, the interaction between the spins is almost negligible so that their orientations can be considered statistically independent. The probability that any given moment points up is, therefore, unaffected by whether any other moment in the system points up or down.

Among the N magnetic moments of the spin system, let us denote by n the number of magnetic moments which point up and by n' the number of magnetic moments which point down. Of course,

$$n + n' = N \quad (7)$$

so that $n' = N - n$. Consider then the spin systems in the statistical ensemble. The number n of moments pointing up is not the same in each system, but can assume any of the possible values $n = 0, 1, 2, \dots, N$. The question of interest is then the following: For each pos-

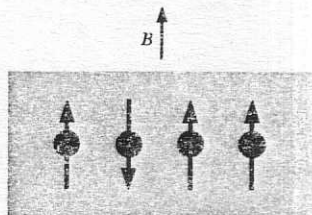


Fig. 2.5 A system consisting of N spins $\frac{1}{2}$ in the special case where $N = 4$. Each arrow indicates the direction of the magnetic moment of a spin. The external magnetic field is denoted by $\mathbf{B}$.

† We shall consider p and q as given quantities determined by experiment. In Chap. 4 we shall learn how to calculate p and q for any value of B if the spin system is known to be at a given temperature.

1	2	3	4	n	n'	$C_N(n)$
U	U	U	U	4	0	1
U	U	U	D	3	1	4
U	U	D	U	3	1	
U	D	U	U	3	1	
D	U	U	U	3	1	
U	U	D	D	2	2	6
U	D	U	D	2	2	
U	D	D	U	2	2	
D	U	U	D	2	2	
D	U	D	U	2	2	
D	D	U	U	2	2	
U	D	D	D	1	3	4
D	U	D	D	1	3	
D	D	U	D	1	3	
D	D	D	U	1	3	
D	D	D	D	0	4	1

Table 2.1. A table T which lists all the possible orientations of N magnetic moments in the special case where $N = 4$. The letter U indicates a moment which points up, the letter D a moment which points down. The number of moments pointing up is denoted by n , the number pointing down by n' . The number $C_N(n)$ of possible configurations in which n of the N moments point up is indicated in the last column. (Note that this table is analogous to Table 1.1.)

sible value of n , what is the probability $P(n)$ that n of the N magnetic moments point up?

The problem of finding the probability $P(n)$ is readily solved by the following argument. The probability that any one magnetic moment points up is p , the probability that it points down is $q = 1 - p$. Since all the magnetic moments are statistically independent, the general †

$$\left[\begin{array}{l} \text{the probability of occurrence of} \\ \text{one particular configuration} \\ \text{where } n \text{ moments point up and} \\ \text{the remaining } n' \text{ point down} \end{array} \right] = \underbrace{pp \cdots p}_{n \text{ factors}} \underbrace{qq \cdots q}_{n' \text{ factors}} = p^n q^{n'}. \quad (8)$$

But a situation where any n moments point up can ordinarily be achieved in many alternative ways, as is illustrated in Table 2.1. Let us, therefore, introduce the notation

$$C_N(n) \equiv \text{the number of distinct configurations of } N \text{ moments} \\ \text{where any } n \text{ of these moments point up (and the} \quad (9) \\ \text{remaining } n' \text{ point down).}^\dagger$$

The desired probability $P(n)$ that n of the N moments point up is equal to the probability that either the first, or the second, ..., or the last of the $C_N(n)$ alternative possibilities is realized. In accordance with the general relation (4), the probability $P(n)$ is thus obtained by summing the probability (8) over the $C_N(n)$ possible configurations having n moments pointing up, i.e., by multiplying the probability (8) by $C_N(n)$. Thus we get

$$P(n) = C_N(n) p^n q^{N-n} \quad (10)$$

where we have put $n' = N - n$.

It only remains to calculate the number of configurations $C_N(n)$ in the general case of arbitrary N and n . Suppose then that we consider a table T , similar to Table 2.1, in which we list all the possible configurations of the N moments and designate by U each moment which points up and by D each moment which points down. The number $C_N(n)$ is then the number of entries in which the letter U appears n times. To examine how many such entries there are, let us consider n distinct moments which point up and label them by the letters $U_1, U_2, \dots, U_n$. In how many ways can we list these in a table T' (such as Table 2.2 shown for the special case where $N = 4$ and $n = 2$)?

† The number $C_N(n)$ is sometimes called the number of combinations of N objects taken n at a time.

The letter U_1 can be listed in a line of the table in any one of N different places;

for each possible location of U_1 , the letter U_2 can then be listed in any one of the remaining $(N - 1)$ places;

for each possible location of U_1 and U_2 , the letter U_3 can then be listed in any one of the remaining $(N - 2)$ places;

.....
for each possible location of $U_1, U_2, \dots, U_{n-1}$, the last letter U_n can then be listed in any one of the remaining $(N - n + 1)$ places.

- The possible number $J_N(n)$ of distinct entries in the table T' is then obtained by multiplying together the number of possible locations of the letters $U_1, U_2, \dots, U_n$; i.e.,

$$J_N(n) = N(N - 1)(N - 2) \cdots (N - n + 1). \quad (11)$$

This can be written more compactly in terms of factorials. Thus†

$$\begin{aligned} J_N(n) &= \frac{N(N - 1)(N - 2) \cdots (N - n + 1)(N - n) \cdots (1)}{(N - n) \cdots (1)} \\ &= \frac{N!}{(N - n)!}. \end{aligned} \quad (12)$$

- The symbols $U_1, U_2, \dots, U_n$ were considered distinct in the preceding enumeration, while the subscripts are really irrelevant since all moments pointing up are equivalent; i.e., any U_i denotes a moment pointing up, irrespective of i . Those entries in the table T' which differ only by a permutation of the subscripts correspond, therefore, to physically indistinguishable situations (see, for example, Table 2.2). Since the possible number of permutations of the n subscripts is given by $n!$, the table T' contains $n!$ times too many entries if only distinct nonequivalent entries are to be considered.‡ The desired number $C_N(n)$ of distinct configurations of up and down moments is then given by dividing $J_N(n)$ by $n!$ so that

$$C_N(n) = \frac{J_N(n)}{n!} = \frac{N!}{n!(N - n)!}. \quad (13)$$

1	2	3	4	
U_1	U_2			a
U_1		U_2		b
U_1			U_2	c
U_2	U_1			a
	U_1	U_2		d
	U_1		U_2	e
U_2		U_1		b
	U_2	U_1		d
		U_1	U_2	f
U_2			U_1	c
	U_2		U_1	e
		U_2	U_1	f

Table 2.2. A table T' which lists all the possible arrangements of $N = 4$ identical moments, $n = 2$ of which point up. For ease of enumeration, these two moments have been labeled by U_1 and U_2 , respectively, although they are physically indistinguishable. Hence entries which differ only in subscripts are equivalent. Such equivalent entries are indicated by identical letters in the last column. The table contains thus $n! = 2$ times too many entries if one is interested only in entries which are physically distinct.

† By definition, $N! \equiv N(N - 1)(N - 2) \cdots (1)$; in addition, $0! \equiv 1$.

‡ The first subscript can assume any of n possible values, the second subscript any of the remaining $(n - 1)$ possible values, ..., and the n th subscript the one remaining value. Hence the subscripts can be arranged in $n(n - 1) \cdots (1) \equiv n!$ possible ways.

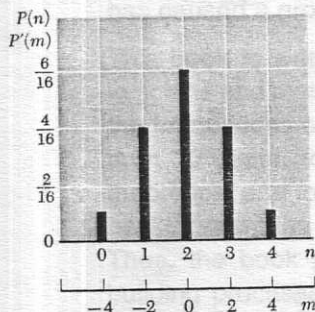


Fig. 2.6 Binomial distribution for $N = 4$ magnetic moments when $p = q = \frac{1}{2}$. The graph shows the probability $P(n)$ that n moments point up, or equivalently, the probability $P'(m)$ that the total magnetic moment in the up direction is equal to m (when measured in units of μ_0).

The desired probability (10) then becomes

$$P(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad (14)$$

or, in more symmetrical form,

$$P(n) = \frac{N!}{n!n'} p^n q^{n'}, \quad \text{where } n' \equiv N - n. \quad (15)$$

In the special case

$$\text{when } p = q = \frac{1}{2}, \quad P(n) = \frac{N!}{n!n'} \left(\frac{1}{2}\right)^N. \quad (16)$$

For a given number N , the probability $P(n)$ is a function of n and is called the *binomial distribution*.

Remark

In expanding a binomial of the form $(p + q)^N$, the coefficient of the term $p^n q^{N-n}$ is simply equal to the number $C_N(n)$ of possible terms which involve the factor p exactly n times and the factor q exactly $(N - n)$ times. Hence one obtains the purely mathematical result known as the *binomial theorem*

$$(p + q)^N = \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n q^{N-n}. \quad (17)$$

Comparison with (14) shows that each

term on the right is precisely the probability $P(n)$. This is the reason for the name "binomial distribution." Incidentally, since $p + q = 1$ when p and q are the probabilities of interest to us, Eq. (17) is equivalent to

$$1 = \sum_{n=0}^N P(n).$$

This verifies that the sum of the probabilities for all possible values of n is properly equal to unity, as required by the normalization condition (3).

Discussion

- † To examine the dependence of $P(n)$ on n , we investigate first the behavior of the coefficient $C_N(n)$ given by (13). Note first that $C_N(n)$ is symmetric under interchange of n and $N - n = n'$. Thus

$$C_N(n') = C_N(n). \quad (18)$$

In addition,

$$C_N(0) = C_N(N) = 1. \quad (19)$$

We also note that

$$\frac{C_N(n+1)}{C_N(n)} = \frac{n!(N-n)!}{(n+1)!(N-n-1)!} = \frac{N-n}{n+1}. \quad (20)$$

Starting with $n = 0$, the ratio of successive coefficients $C_N(n)$ is thus initially large, of the order of N ; it then decreases monotonically with n , staying larger than (or at most becoming equal to) unity as long as $n < \frac{1}{2}N$, and becoming smaller than unity for $n \geq \frac{1}{2}N$. This behavior, combined with (19), shows that $C_N(n)$ has a maximum near $n = \frac{1}{2}N$ and that its value there is very large compared to unity, to the extent that N is large.

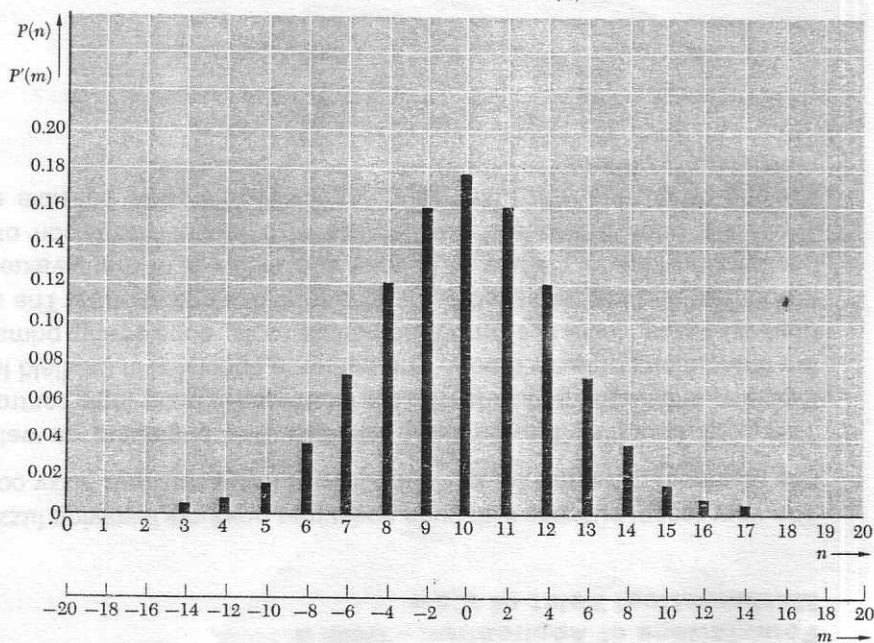
The behavior of the probability $P(n)$ is then apparent. By (16) it follows that,

$$\text{if } p = q = \frac{1}{2}, \quad P(n') = P(n). \quad (21)$$

This result must, of course, be true by symmetry since there is no preferred spatial orientation if $p = q$ (i.e., in the absence of an applied magnetic field $\mathbf{B}$). In this case the probability $P(n)$ has a maximum† near $n = \frac{1}{2}N$. On the other hand, if $p > q$, the coefficient $C_N(n)$ still tends to produce a maximum of $P(n)$, but this maximum is now shifted to a value $n > \frac{1}{2}N$. Figures 2.6 and 2.7 illustrate the behavior of the probability $P(n)$ in some simple cases.

Fig. 2.7 Binominal distribution for $N = 20$ magnetic moments when $p = q = \frac{1}{2}$. The graph shows the probability $P(n)$ that n moments point up, or equivalently, the probability $P'(m)$ that the total magnetic moment in the up direction is equal to m (when measured in units of μ_0).

† The maximum is at $\frac{1}{2}N$ when N is even, and straddles this value otherwise.



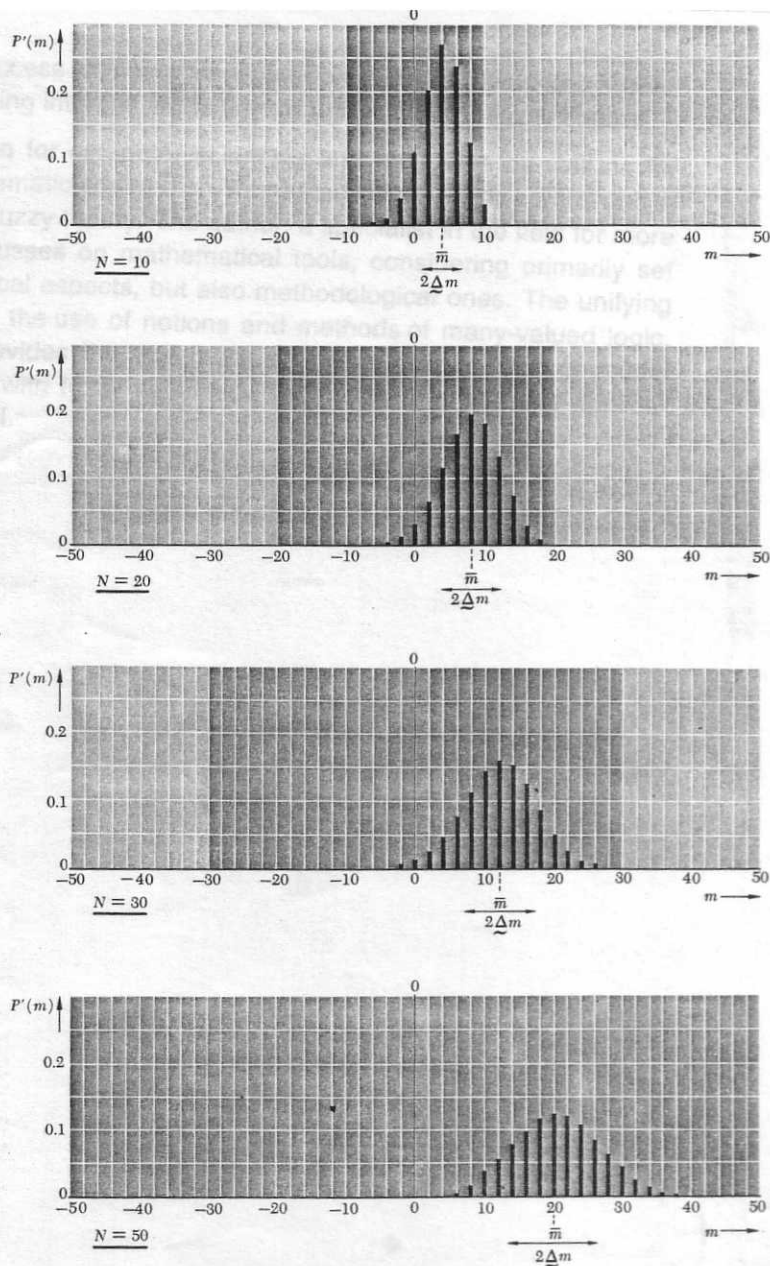


Fig. 2.8 The probability $P'(m)$ that the total magnetic moment of a system of N spins $\frac{1}{2}$ is equal to m (measured in units of μ_0). Because of the presence of an applied magnetic field, $p = 0.7$ and $q = 0.3$. The graphs show $P'(m)$ in four different cases corresponding to $N = 10$, $N = 20$, $N = 30$, and $N = 50$.

The total magnetic moment of a spin system is a quantity which is readily measured experimentally. Let us denote by M the total magnetic moment along the "up" direction. Since M is simply equal to the algebraic sum of the components along the "up" direction of the magnetic moments of all the N spins, it follows that

$$M = n\mu_0 - n'\mu_0 = (n - n')\mu_0$$

$$\text{or} \quad M = m\mu_0 \quad (22)$$

$$\text{where} \quad m \equiv n - n' \quad (23)$$

and where μ_0 denotes the magnitude of the magnetic moment of a spin. By (22), $m = M/\mu_0$ is simply the total magnetic moment measured in units of μ_0 . The relation (23) can also be written as

$$m = n - n' = n - (N - n) = 2n - N. \quad (24)$$

This shows, incidentally, that the possible values of m must be odd if N is odd, and must be even when N is even. According to (24), a definite value of n corresponds to a unique value of m , and conversely

$$n = \frac{1}{2}(N + m). \quad (25)$$

The probability $P'(m)$ that m assumes a certain value must then be the same as the probability $P(n)$ that n assumes the corresponding value given by (25). Thus

$$P'(m) = P\left(\frac{N + m}{2}\right). \quad (26)$$

This expression gives the probability of occurrence of any possible value of the total magnetic moment of the spin system. In the special case where $p = q = \frac{1}{2}$, the expressions (16) and (26) yield thus explicitly

$$P'(m) = \frac{N!}{\left(\frac{N + m}{2}\right)! \left(\frac{N - m}{2}\right)!} \left(\frac{1}{2}\right)^N;$$

the most probable situation is then clearly at (or near) $m = 0$ where $M = 0$.

Generality of the binomial distribution

Although our discussion has dealt with the specific problem of a system of spins, it can be phrased more abstractly. Thus we have really solved the following general problem: Given N events which are statistically independent. Suppose that each such event occurs with a probability p ; the probability that it does not occur is thus given by

- $q = 1 - p$. What then is the probability $P(n)$ that any n out of these N events do occur (while the remaining $n = N - n$ events do not occur)? This question is immediately answered by the binomial distribution (14). Indeed, in our specific example of a system of N independent spins, the occurrence of an event is merely represented by a spin pointing up, while the nonoccurrence of an event is represented by a spin not pointing up, i.e., pointing down.

Some further examples will serve to illustrate some common problems which are immediately answered by the binomial distribution.

Ideal gas of N molecules

Consider an ideal gas of N molecules enclosed in a box of volume V_0 . Since the molecules of an ideal gas interact with each other to an almost negligible extent, their motion is statistically independent.

- + Suppose that the box is imagined to be subdivided into two parts of respective volumes V and V' , where

$$V + V' = V_0. \quad (27)$$

- Consider an ensemble of many such boxes of gas. Let p denote the probability that a given molecule is found in the volume V , and q the probability that it is found in the remaining volume V' .
- If the gas is in equilibrium, each molecule tends to be uniformly distributed

throughout the box so that

$$p = \frac{V}{V_0} \quad \text{and} \quad q = \frac{V'}{V_0}. \quad (28)$$

Thus $p + q = 1$, as in (6). What then is the probability $P(n)$ in the ensemble that n of the N molecules are found in the volume V (while the remaining $n' = N - n$ molecules are found in the volume V')? The answer is given by the binomial distribution (14). In particular, if $V = V'$ so that $p = q = \frac{1}{2}$, we have thus solved explicitly the problem of Sec. 1.1 where we wanted to find the probability that n out of N molecules are found in the left half of a box.

Tossing of coins or dice

Consider the tossing of a set of N coins. The behavior of these coins can be considered statistically independent. Let p denote the probability that any given coin shows a head and q the probability that it shows a tail. By symmetry, we can assume that $p = q = \frac{1}{2}$. What then is the probability $P(n)$ that n of the N coins show heads?

The tossing of a set of N dice is similar. These again can be considered statisti-

cally independent. Let p denote the probability that any given die lands with a "6" uppermost and $q = 1 - p$ the probability that it does not. Since a die has 6 faces, we can assume by symmetry that $p = \frac{1}{6}$ while $q = 1 - p = \frac{5}{6}$. What then is the probability $P(n)$ that n of the N dice show a "6" uppermost? This question again is answered by the binomial distribution (14).

2.4 Mean Values

Suppose that a variable u of some system can assume any of α possible distinct values

$$u_1, u_2, \dots, u_\alpha$$

with respective probabilities

$$P_1, P_2, \dots, P_\alpha.$$

This means that, in an ensemble of $\mathcal{N}$ similar systems (where $\mathcal{N} \rightarrow \infty$), the variable u assumes the particular value u_r in a number $\mathcal{N}_r = \mathcal{N}P_r$ of these systems.

The specification of the probabilities P_r for all the α possible values u_r of the variable u constitutes the most complete statistical description of the system. It is, however, also convenient to define parameters that characterize in a less detailed way the distribution of the possible values of u in the ensemble. These parameters are certain *mean* (or *average*) values. The notion is quite familiar. For example, the result of an examination given to a group of students can be described most completely (if one does not wish to name individual students) by specifying the number of students receiving each of the possible grades given on this examination. But the result can also be characterized in a less detailed manner by computing the mean grade of the students. This is conventionally done by multiplying each possible grade by the number of students receiving this grade, adding all the resultant products, and then dividing this sum by the total number of students. In a similar manner, the mean value of u in the ensemble is defined by multiplying each possible value u_r by the number $\mathcal{N}_r$ of systems in the ensemble which exhibit this value, adding the resultant products for all α possible values of the variable u , and then dividing this sum by the total number $\mathcal{N}$ of systems in the ensemble. The *mean value* of u (or *ensemble average* of u), which we shall denote by $\bar{u}$, is thus defined by

$$\bar{u} \equiv \frac{\mathcal{N}_1 u_1 + \mathcal{N}_2 u_2 + \dots + \mathcal{N}_\alpha u_\alpha}{\mathcal{N}} = \frac{\sum_{r=1}^{\alpha} \mathcal{N}_r u_r}{\mathcal{N}}. \quad (29)$$

But, since $\mathcal{N}_r/\mathcal{N} \equiv P_r$ is the probability of occurrence of the value u_r , the definition (29) becomes simply†

$$\bar{u} \equiv \sum_{r=1}^{\alpha} P_r u_r. \quad (30)$$

Similarly, if $f(u)$ is any function of u , the *mean value* (or *ensemble average*) of f is defined by the expression

$$\bar{f}(u) \equiv \sum_{r=1}^{\alpha} P_r f(u_r). \quad (31)$$

This definition implies that mean values have some very simple properties. For example, if $f(u)$ and $g(u)$ are any two functions of u ,

$$\overline{f+g} \equiv \sum_{r=1}^{\alpha} P_r [f(u_r) + g(u_r)] = \sum_{r=1}^{\alpha} P_r f(u_r) + \sum_{r=1}^{\alpha} P_r g(u_r)$$

or

$$\overline{f+g} = \bar{f} + \bar{g}. \quad (32)$$

This result shows quite generally that the mean value of a sum of terms is equal to the sum of the mean values of these terms. Thus the successive operations of performing a sum and taking an average yield the same result irrespective of the order in which they are carried out.‡

Similarly, if c is any constant,

$$\overline{cf} = \sum_{r=1}^{\alpha} P_r [cf(u_r)] = c \sum_{r=1}^{\alpha} P_r f(u_r)$$

or

$$\overline{cf} = c\bar{f}. \quad (33)$$

Thus the operations of multiplying by a constant and taking an average can also be carried out in either order without affecting the result. If $f = 1$, the relation (33) makes the obvious assertion that the mean value of a constant is simply equal to this constant.

† The mean value $\bar{u}$ is time-dependent if the ensemble is time-dependent, i.e., if some of the probabilities P_r depend on time. Note also that the *mean value* or *ensemble average* $\bar{u}$ is an average over all systems in the ensemble at a particular time. It is ordinarily different from the time average defined in (1.6) for a single system, except in the special case of time-independent ensembles where the time average is taken over a very long interval of time.

‡ In mathematical language one would say that these operations “commute.”

Example

Consider a system of 4 spins when $p = q = \frac{1}{2}$. The number of moments pointing up can then be $n = 0, 1, 2, 3, 4$. These numbers occur with probabilities $P(n)$ which follow immediately from (16) and which were already calculated very simply in (1.4a). As indicated in Fig. 2.6, these probabilities are respectively,

$$P(n) = \frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16}.$$

The mean number of moments pointing up is then

$$\begin{aligned}\bar{n} &= \sum_{n=0}^4 P(n)n \\ &= \left(\frac{1}{16} \times 0\right) + \left(\frac{4}{16} \times 1\right) + \left(\frac{6}{16} \times 2\right) \\ &\quad + \left(\frac{4}{16} \times 3\right) + \left(\frac{1}{16} \times 4\right) \\ &= 2.\end{aligned}$$

Note that this result is simply equal to $Np = 4 \times \frac{1}{2}$.

Since $p = q$, there is no preferred direction in space. The mean number of moments pointing down must thus be equal to the mean number of moments

pointing up, i.e.,

$$n' = \bar{n} = 2.$$

This result also follows from the relation (32) which allows us to write

$$\bar{n}' = \overline{N - n} = \bar{N} - \bar{n} = 4 - 2 = 2.$$

Since there is no preferred direction in space, the mean magnetic moment must clearly vanish. Indeed, one has

$$\bar{m} = \overline{n - n'} = \bar{n} - \bar{n}' = 2 - 2 = 0.$$

The value of $\bar{m}$ could also be computed directly, of course, by using the probability $P'(m)$ that m assumes its possible values $m = -4, -2, 0, 2, 4$. Thus one has, by definition,

$$\begin{aligned}\bar{m} &= \sum_m P'(m)m \\ &= \left[\frac{1}{16} \times (-4)\right] + \left[\frac{4}{16} \times (-2)\right] \\ &\quad + \left[\frac{6}{16} \times 0\right] + \left[\frac{4}{16} \times 2\right] + \left[\frac{1}{16} \times 4\right] \\ &= 0.\end{aligned}$$

•• One last property of mean values is often of importance. Suppose that one is dealing with two variables, u and v , which can assume the values

$$u_1, u_2, \dots, u_\alpha$$

and

$$v_1, v_2, \dots, v_\beta,$$

respectively. Let us denote by P_r the probability that u assumes the value u_r , and by P_s the probability that v assumes the value v_s . If the probability that u assumes any of its values is independent of the value assumed by v (i.e., if the variables u and v are statistically independent), then the joint probability P_{rs} that u assumes the value u_r and that v assumes the value v_s is, by (5), simply equal to

$$P_{rs} = P_r P_s. \quad (34)$$

+ Suppose now that $f(u)$ is any function of u while $g(v)$ is any function of v . Then the mean value of the product fg is, by the definition (31), quite generally given by

$$\overline{f(u)g(v)} \equiv \sum_{r=1}^{\alpha} \sum_{s=1}^{\beta} P_{rs} f(u_r) g(v_s) \quad (35)$$

where the summation is over all possible values u_r and v_s of the variables. If the variables are statistically independent so that (34) is valid, Eq. (35) becomes

$$\begin{aligned}\bar{fg} &= \sum_r \sum_s P_r P_s f(u_r) g(v_s) \\ &= \sum_r \sum_s [P_r f(u_r)] [P_s g(v_s)] \\ &= \left[\sum_r P_r f(u_r) \right] \left[\sum_s P_s g(v_s) \right].\end{aligned}$$

But the first of the factors on the right is simply the mean value of f , while the second is the mean value of g . Hence we arrive at the result that

$$\begin{array}{|l} \dagger \\ \hline \text{if } u \text{ and } v \text{ are statistically independent,} \\ \hline \bar{fg} = \bar{f} \bar{g}; \end{array} \quad (36)$$

i.e., the average of a product is then simply equal to the product of the averages.

Dispersion

Suppose that a variable u assumes its possible values u_r with respective probabilities P_r . Some general features of the probability distribution can then be characterized by a few useful parameters. One of these is simply the mean value of u itself, i.e., the quantity $\bar{u}$ defined in (30). This parameter indicates the central value of u about which the various values u_r are distributed. It is then often convenient to measure the possible values of u with respect to their mean value $\bar{u}$ by putting

$$\Delta u \equiv u - \bar{u} \quad (37)$$

where Δu is the deviation of u from the mean value $\bar{u}$. Note that the mean value of this deviation vanishes. Indeed, using the property (32),

$$\overline{\Delta u} = (\overline{u - \bar{u}}) = \bar{u} - \bar{u} = 0. \quad (38)$$

It is also useful to define a parameter which measures the extent of the spread of the possible values of u about their mean value $\bar{u}$. The mean value of Δu itself does not provide such a measure since Δu is, on the average, as often positive as negative so that its mean value vanishes in accordance with (38). On the other hand, the quantity $(\Delta u)^2$ can never be negative. Its mean value, defined by

$$\overline{(\Delta u)^2} \equiv \sum_{r=1}^{\alpha} P_r (\Delta u_r)^2 \equiv \sum_{r=1}^{\alpha} P_r (u_r - \bar{u})^2 \quad (39)$$

is called the *dispersion* (or *variance*) of u and also can never be negative since each term in the sum of (39) is nonnegative.† Thus

$$(\overline{\Delta u})^2 \geq 0. \quad (40)$$

† The dispersion can only vanish if *all* occurring values of u_r are equal to $\bar{u}$; it becomes increasingly large to the extent that these values have appreciable probability of occurring far from $\bar{u}$. The dispersion does, therefore, provide a convenient measure of the amount of scatter of the values assumed by u .

The dispersion $(\overline{\Delta u})^2$ is a quantity which has the dimensions of the square of u . A linear measure of the spread of possible values of u is provided by the square root of the dispersion, i.e., by the quantity

$$\Delta u \equiv [(\overline{\Delta u})^2]^{1/2} \quad (41)$$

which has the same dimensions as u itself and which is called the *standard deviation* of u . The definition (39) shows that even a few values of u occurring with appreciable probability far from $\bar{u}$ would make a large contribution to Δu . Most values of u must therefore occur within a range of the order of Δu surrounding the mean value $\bar{u}$.

Example

Let us return to the previous example of four spins when $p = q = \frac{1}{2}$. Since $\bar{n} = 2$, the dispersion of n is, by definition,

$$\begin{aligned} (\overline{\Delta n})^2 &\equiv \sum_n P(n)(n - 2)^2 \\ &= [\tfrac{1}{16} \times (-2)^2] + [\tfrac{1}{4} \times (-1)^2] \\ &\quad + [\tfrac{6}{16} \times (0)^2] + [\tfrac{1}{16} \times (1)^2] \\ &\quad + [\tfrac{1}{16} \times (2)^2] \\ &= 1. \end{aligned}$$

Hence the standard deviation of n is

$$\Delta n = \sqrt{1} = 1.$$

Similarly one can calculate the dispersion of the magnetic moment. Since $\bar{m} = 0$, one has, by definition,

$$\begin{aligned} (\overline{\Delta m})^2 &\equiv \sum_m P'(m)(m - 0)^2 \\ &= [\tfrac{1}{16} \times (-4)^2] + [\tfrac{1}{4} \times (-2)^2] \\ &\quad + [\tfrac{6}{16} \times (0)^2] + [\tfrac{1}{16} \times (2)^2] \\ &\quad + [\tfrac{1}{16} \times (4)^2] \\ &= 4 \end{aligned}$$

so that

$$\Delta m = \sqrt{4} = 2.$$

Let us verify that the preceding results are consistent. Since $\bar{m} = 0$ while $\bar{n} = \bar{n}' = 2$, one has for all values of m or n

$$\begin{aligned} \Delta m &= m - n - n' = n - (4 - n) \\ &= 2n - 4 = 2(n - 2) \end{aligned}$$

$$\text{or} \quad \Delta m = 2(n - \bar{n}) = 2 \Delta n.$$

$$\text{Hence } (\overline{\Delta m})^2 = 4(\overline{\Delta n})^2,$$

in agreement with what we found by explicit calculation.

† Note that $(\overline{\Delta u})^2$ is different from the quantity $\overline{(\Delta u)^2}$; i.e., it makes a great deal of difference whether one squares first and then takes the average, or whether one carries out these operations in the reverse order.

- † A knowledge of the probabilities P_r for all values u_r gives complete statistical information about the distribution of the values of u in the ensemble. On the other hand, a knowledge of a few mean values such as $\bar{u}$ and $(\Delta u)^2$ provides only a partial knowledge of the characteristics of this distribution and is not sufficient to determine the probabilities P_r unambiguously. Such mean values, however, may often be calculated very simply without an explicit knowledge of the probabilities, even in cases where the actual computation of these probabilities would be a difficult task. We shall illustrate these remarks in the next section.

2.5 Calculation of Mean Values for a Spin System

- † Consider an ideal system of N spins $\frac{1}{2}$. The fact that these spins are statistically independent allows us to calculate various mean values quite simply under very general conditions. The calculation can be accomplished without the need of computing any probabilities such as the probability $P(n)$ obtained in (14).

Let us start from scratch then to investigate a physically interesting quantity of this spin system, its total magnetic moment M along the up direction. Let us denote by μ_i the component along the up direction of the i th spin. The total magnetic moment M is then simply equal to the sum of the magnetic moments of all the spins so that

$$M = \mu_1 + \mu_2 + \cdots + \mu_N$$

or, in more compact notation,

$$M = \sum_{i=1}^N \mu_i. \quad (42)$$

We should like to calculate the mean value and the dispersion of this total magnetic moment.

- (1) To compute the mean value of M , we need only take the mean values of both sides of (42). The general property (32), which allows us to interchange the order of averaging and summing, immediately yields the result

$$\bar{M} = \sum_{i=1}^N \bar{\mu}_i = \sum_{i=1}^N \bar{\mu}_1. \quad (43)$$

But the probability that any magnetic moment has a given orientation (either up or down) is the same for each moment; hence the mean magnetic moment is the same for each spin (i.e., $\bar{\mu}_1 = \bar{\mu}_2 = \cdots = \bar{\mu}_N$) and can be denoted simply by $\bar{\mu}$. The sum in (43) consists, therefore, of N equal terms so that (43) becomes simply

$$\begin{aligned} \bar{n} &= \sum_{n=0}^N P(n) n = \\ &= \sum_{n=0}^N \frac{N!}{n!(N-n)!} n p^n q^{N-n} \\ &= \sum_{n=0}^N \frac{N!}{(n-1)!(N-n)!} p^n q^{N-n} \\ &= \sum_{n=1}^N \frac{N!}{(n-1)!(N-n)!} p^n q^{N-n} \\ &= p \sum_{n=1}^N \frac{N!}{(n-1)!(N-n)!} p^{n-1} q^{N-n} \\ &= p \sum_{n=0}^{N-1} \frac{N!}{n!(N-n)!} p^n q^{N-n} \\ &= p \sum_{n=0}^{N-1} P(n) = p \sum_{n=0}^N P(n) = p \cdot 1 = p \end{aligned}$$

$$\bar{M} = N\bar{\mu}. \quad (44)$$

This result is almost self-evident; it asserts merely that the mean total magnetic moment of N spins is N times as large as the mean moment of one spin.

(2) Now let us calculate the dispersion of M , i.e., the quantity $\overline{(\Delta M)^2}$ where

$$\Delta M \equiv M - \bar{M}. \quad (45)$$

Subtraction of (43) from (42) yields

$$M - \bar{M} = \sum_{i=1}^N (\mu_i - \bar{\mu})$$

or

$$\Delta M = \sum_{i=1}^N \Delta \mu_i, \quad (46)$$

where

$$\Delta \mu_i \equiv \mu_i - \bar{\mu}. \quad (47)$$

To find $(\Delta M)^2$, we need only multiply the sum in (46) by itself. Thus

$$\begin{aligned} (\Delta M)^2 &= (\Delta \mu_1 + \Delta \mu_2 + \cdots + \Delta \mu_N) (\Delta \mu_1 + \Delta \mu_2 + \cdots + \Delta \mu_N) \\ &= [(\Delta \mu_1)^2 + (\Delta \mu_2)^2 + (\Delta \mu_3)^2 + \cdots + (\Delta \mu_N)^2] \\ &\quad + [\Delta \mu_1 \Delta \mu_2 + \Delta \mu_1 \Delta \mu_3 + \cdots + \Delta \mu_1 \Delta \mu_N \\ &\quad + \Delta \mu_2 \Delta \mu_1 + \Delta \mu_2 \Delta \mu_3 + \cdots + \Delta \mu_N \Delta \mu_{N-1}] \end{aligned}$$

or

$$(\Delta M)^2 = \sum_{i=1}^N (\Delta \mu_i)^2 + \sum_{i=1}^N \sum_{j=1, j \neq i}^N (\Delta \mu_i)(\Delta \mu_j). \quad (48)$$

The first term on the right represents all the squared terms arising from terms in the sum (46) which are multiplied by themselves; the second term represents all the cross terms arising from products of different terms in the sum (46). Taking the mean value of (48) and using again the property (32) which allows us to interchange the order of averaging and summing, we then obtain

$$\overline{(\Delta M)^2} = \sum_{i=1}^N \overline{(\Delta \mu_i)^2} + \sum_{i=1}^N \sum_{j=1, j \neq i}^N \overline{(\Delta \mu_i)(\Delta \mu_j)}. \quad (49)$$

All products in the second sum, where $i \neq j$, refer to different spins. But since different spins are statistically independent, the property (36) implies that the mean value of each such product is simply equal to the product of the mean values of its factors. Thus,

for $i \neq j$, $\overline{(\Delta \mu_i)(\Delta \mu_j)} = (\overline{\Delta \mu_i})(\overline{\Delta \mu_j}) = 0$ (50)
since $\overline{\Delta \mu_i} = \bar{\mu}_i - \bar{\mu} = 0$.

$$\begin{aligned} \Delta M &= \sum_{n=0}^N P(n) (n - \bar{n}) \text{ but} \\ (n - \bar{n})^2 &= n^2 - 2n\bar{n} + \bar{n}^2 \Rightarrow \\ \Delta M &= \sum P(n) n^2 - 2\bar{n} \sum_{n=0}^N n P(n) + \bar{n}^2 \\ &= \sum_{n=0}^N n^2 P(n) - \bar{n}^2, \\ \sum_{n=0}^N n^2 P(n) &= \sum_{n=0}^N n^2 p^n q^{N-n} \frac{N!}{n!(N-n)!} \\ \text{but } n^2 p^n &= (p \frac{\partial}{\partial p})^2 p^n \Rightarrow \\ \overline{(n^2)} &= \sum_{n=0}^N (p \frac{\partial}{\partial p})^2 p^n q^{N-n} \frac{N!}{n!(N-n)!} \\ &= \overline{(n^2)} = 2(p \frac{\partial}{\partial p})^2 (p+q)^N, \\ p \frac{\partial}{\partial p} (p+q)^N &= NP(p+q)^{N-1} \\ \Rightarrow \overline{(n^2)} &= p \frac{\partial}{\partial p} (NP(p+q)^{N-1}) = \\ NP [N-1] p (p+q)^{N-2} + (p+q)^{N-1} &= \\ NP [N-1] p + NP(p+q) &= NP(Np+q) \\ \Rightarrow \Delta M &= N^2 p^2 + NPq - N\bar{n}^2 \\ \Rightarrow \Delta M &= NPq \end{aligned}$$

In short, each cross term in (49) vanishes on the average, being as often positive as negative. Hence (50) reduces simply to a sum of squared terms (none of which can be negative):

$$\overline{(\Delta M)^2} = \sum_{i=1}^N \overline{(\Delta \mu_i)^2} \quad (51)$$

* The argument now becomes identical to that following Eq. (43). The probability that any moment has any given orientation is the same for each moment; hence the dispersion $(\Delta \mu_i)^2$ is the same for each spin [i.e., $\overline{(\Delta \mu_1)^2} = \overline{(\Delta \mu_2)^2} = \dots = \overline{(\Delta \mu_N)^2}$] and can be denoted simply by $\overline{(\Delta \mu)^2}$. The sum in (51) consists thus of N equal terms and reduces simply to

$$\boxed{\overline{(\Delta M)^2} = N \overline{(\Delta \mu)^2}} \quad (52)$$

* This relation asserts that the dispersion of the total magnetic moment is merely N times as large as the dispersion of the magnetic moment of an individual spin. Correspondingly (52) also implies that

$$\boxed{\Delta M = \sqrt{N} \Delta \mu} \quad (53)$$

where $\Delta M \equiv [\overline{(\Delta M)^2}]^{1/2}$ and $\Delta \mu \equiv [\overline{(\Delta \mu)^2}]^{1/2}$

are, in accordance with the general definition (41), the standard deviations of the total magnetic moment and of the magnetic moment per spin, respectively.

* The relations (44) and (53) show explicitly how $\bar{M}$ and ΔM depend on the total number N of spins in the system. When $\bar{\mu} \neq 0$, the mean total magnetic moment $\bar{M}$ increases proportionately to N . The standard deviation ΔM (which measures the width of the distribution of values of M about their mean value $\bar{M}$) also increases as N increases, but does so only proportionately to $N^{1/2}$. Hence the relative magnitude of ΔM compared to $\bar{M}$ decreases proportionately to $N^{-1/2}$; indeed (44) and (53) imply that

$$\text{for } \bar{\mu} \neq 0, \quad \frac{\Delta M}{\bar{M}} = \frac{1}{\sqrt{N}} \left(\frac{\Delta \mu}{\bar{\mu}} \right). \quad (54)$$

Figure 2.8 illustrates these characteristic trends.

(3) Note that the results (44) and (53) are very general. They depend only on the additive relation (43) and on the fact that the spins are statistically independent. All our considerations would thus remain equally valid even if the component μ_i of each magnetic moment

* could assume many possible values. (This would be the case if the spin of each particle were greater than $\frac{1}{2}$ so that it could exhibit more than two possible orientations in space.)

(1) *System of particles with spin $\frac{1}{2}$*

The foregoing results are readily applied to the familiar special case where each particle has spin $\frac{1}{2}$. As usual we suppose that its magnetic moment then has probability p of pointing up so that $\mu_i = \mu_0$, and probability $q = 1 - p$ of pointing down so that $\mu_i = -\mu_0$. Its mean moment along the up direction is thus

$$\bar{\mu} \equiv p\mu_0 + q(-\mu_0) = (p - q)\mu_0 = (2p - 1)\mu_0. \quad (55)$$

As a check we note that, in the symmetrical case where $p = q$, $\bar{\mu} = 0$ as expected.

2 The dispersion of the magnetic moment of a spin is given by

$$\overline{(\Delta\mu)^2} \equiv \overline{(\mu - \bar{\mu})^2} \equiv p(\mu_0 - \bar{\mu})^2 + q(-\mu_0 - \bar{\mu})^2. \quad (56)$$

But $\mu_0 - \bar{\mu} = \mu_0 - (2p - 1)\mu_0 = 2\mu_0(1 - p) = 2\mu_0q$,

and $\mu_0 + \bar{\mu} = \mu_0 + (2p - 1)\mu_0 = 2\mu_0p$.

Thus (56) becomes

$$\overline{(\Delta\mu)^2} = p(2\mu_0q)^2 + q(2\mu_0p)^2 = 4\mu_0^2pq(q + p)$$

$$\text{or} \quad \overline{(\Delta\mu)^2} = 4pq\mu_0^2 \quad (57)$$

since $p + q = 1$.

The relations (44) and (52) yield therefore the results

$$\bar{M} = N(p - q)\mu_0 \quad (58)$$

and

$$\overline{(\Delta M)^2} = 4Npq\mu_0^2. \quad (59)$$

The standard deviation of M is accordingly

$$\underline{\Delta M} = 2\sqrt{Npq}\mu_0. \quad (60)$$

3 If we write $M = m\mu_0$, so that the integer $m = M/\mu_0$ expresses the total magnetic moment in units of μ_0 , the results (58) through (60) can also be expressed in the form

$$\bar{m} = N(p - q) = N(2p - 1), \quad (61)$$

$$\overline{(\Delta m)^2} = 4Npq, \quad (62)$$

$$\underline{\Delta m} = 2\sqrt{Npq}. \quad (63)$$

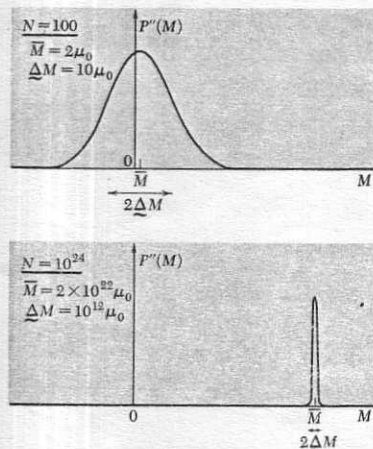


Fig. 2.9 The probability $P''(M)$ that the total magnetic moment of a spin system has a value M when $N = 100$ and when $N = 10^{24}$. The magnetic field is such that $p = 0.51$ and $q = 0.49$. The graphs indicate the envelope curves showing the possible values of $P''(M)$. The two graphs are not drawn to the same scale.

These relations contain an appreciable amount of information about the distribution of the possible values of M or m in the ensemble of spin systems. Thus we know that only those values of m occur with appreciable probability which lie near $\bar{m}$ and do not differ from it by an amount much larger than Δm . Figure 2.8 provides a specific illustration.

Example

Suppose that, in the presence of a certain applied magnetic field $\mathbf{B}$, the magnetic moment of each spin has a probability $p = 0.51$ of pointing parallel to $\mathbf{B}$ and a probability $q = 1 - p = 0.49$ of pointing antiparallel to $\mathbf{B}$. The mean total magnetic moment of a system of N spins is then

$$\bar{M} = 0.02N\mu_0.$$

The standard deviation of its total magnetic moment is given by (60) so that

$$\Delta M = 2\sqrt{Npq}\mu_0 \approx \sqrt{N}\mu_0.$$

$$\text{Hence } \frac{\Delta M}{\bar{M}} \approx \frac{\sqrt{N}\mu_0}{0.02N\mu_0} = \frac{50}{\sqrt{N}}.$$

Consider first a case where the total number of spins is fairly small. For example, suppose that $N = 100$. Then

$$\frac{\Delta M}{\bar{M}} \approx \frac{50}{\sqrt{100}} = 5$$

so that $\Delta M > \bar{M}$. The scatter in the possible values of M is then very pronounced. Indeed, it is quite likely that there occur values of M which differ

widely from $\bar{M}$ and are even of opposite sign. (See Fig. 2.9.)

On the other hand, consider the case of a macroscopic system of spins where N is of the order of Avogadro's number, say $N = 10^{24}$. Then

$$\frac{\Delta M}{\bar{M}} \approx \frac{50}{\sqrt{10^{24}}} = 5 \times 10^{-11}$$

so that $\Delta M \ll \bar{M}$. The scatter in the possible values of M is then very small relative to the mean total magnetic moment. If we set out to measure the total magnetic moment of the system, we would thus almost always measure a value very close to $\bar{M}$. Indeed, unless our method of measurement were sufficiently precise to detect differences of magnetic moment smaller than about one part in 10^{10} , we would virtually always measure a magnetic moment equal to $\bar{M}$ without becoming aware of the existence of fluctuations about this value. This example illustrates concretely the general conclusion that the relative magnitude of fluctuations tends to become very small in a macroscopic system consisting of very many particles.

(ii) Distribution of molecules in an ideal gas

Consider an ideal gas of N molecules contained in a box of volume V_0 .

- We are interested in investigating the number n of molecules found within any specified subvolume V of this box (see Fig. 2.10). If the gas is in equilibrium, then the probability p of finding a molecule in this volume V is simply equal to

$$p = \frac{V}{V_0} \quad (64)$$

as mentioned previously in (28).

It is very easy to calculate the mean value of n and its dispersion. We have already pointed out at the end of Sec. 2.3 that the problem of the ideal gas is analogous to that of the spin system. (Both problems are of the type leading to the binomial distribution.) Hence we can immediately apply the results (61) and (62) to find the desired information about n . Let n' denote the number of molecules in the remaining volume $V_0 - V$ of the box and let $m \equiv n + n'$. As shown in (25), it then follows that

$$n = \frac{1}{2}(N + m). \quad (65)$$

Using the result (61) for $\bar{m}$, we then obtain

$$\bar{n} = \frac{1}{2}(N + \bar{m}) = \frac{1}{2}N(1 + p - q)$$

or
$$\bar{n} = Np \quad (66)$$

since $q = 1 - p$. Furthermore, we obtain from (65) the relation

$$\Delta n \equiv n - \bar{n} = \frac{1}{2}(N + m) - \frac{1}{2}(N + \bar{m}) = \frac{1}{2}(m - \bar{m})$$

or
$$\Delta n = \frac{1}{2} \Delta m.$$

Hence
$$(\Delta n)^2 = \frac{1}{4}(\Delta m)^2$$

and (62) implies that†

$$(\Delta n)^2 = Npq. \quad (67)$$

The standard deviation of n is then

$$\Delta n = \sqrt{Npq} \quad (68)$$

so that
$$\frac{\Delta n}{\bar{n}} = \frac{\sqrt{Npq}}{Np} = \left(\frac{q}{p}\right)^{1/2} \frac{1}{\sqrt{N}}. \quad (69)$$

These relations show again that the standard deviation Δn increases proportionately to $N^{1/2}$. Correspondingly, the relative value $\Delta n/\bar{n}$ of the standard deviation decreases proportionately to $N^{-1/2}$ and thus becomes very small when N is large. These statements are well illustrated by the special case of Chap. 1 where we considered the number n of molecules contained in one half of a box. In this case (64) implies that $p = q = \frac{1}{2}$ so that (66) reduces to the obvious result

$$\bar{n} = \frac{1}{2}N$$

while

$$\frac{\Delta n}{\bar{n}} = \frac{1}{\sqrt{N}}.$$

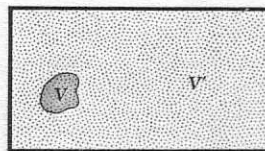


Fig. 2.10 A box of volume V_0 contains N molecules of an ideal gas. At any given time some number n of the molecules are located in the subvolume V , while the remaining number $n' = N - n$ are located in the rest of the volume $V' = V_0 - V$.

† The relations (66) and (67) could also be derived directly by the methods of this section without using the corresponding results for the quantity m (see Prob. 2.14).

These relations put the discussion of fluctuations in Sec. 1.1 on a quantitative basis. The fact that the absolute magnitude of fluctuations, (measured by Δn) increases with increasing N while the relative magnitude of fluctuations (measured by $\Delta n/n$) decreases with increasing N is illustrated explicitly by the graphs of Figs. 1.5 and 1.6 for $N = 4$ and $N = 40$. When the box contains about a mole of gas, N is of the order of Avogadro's number so that $N \sim 10^{24}$. In this case the relative magnitude of fluctuations $\Delta n/n \sim 10^{-12}$ becomes so small as to be almost always negligible.

2.6 Continuous Probability Distributions

Let us consider an ideal spin system consisting of a large number N of spins $\frac{1}{2}$. There are then many possible values of the total magnetic moment of this system. Indeed, by (22) and (24),

$$M = m\mu_0 = (2n - N)\mu_0 \quad (70)$$

so that M can assume any of the $(N + 1)$ possible values

$$M = -N\mu_0, -(N - 2)\mu_0, -(N - 4)\mu_0, \dots, (N - 2)\mu_0, N\mu_0. \quad (71)$$

The probability $P''(M)$ that the total magnetic moment assumes a particular value M is equal to the probability of occurrence of the corresponding value of m or n , i.e., to $P'(m)$ given by (26) or $P(n)$ given by (14).

Thus
$$P''(M) = P'(m) = P(n), \quad (72)$$

where
$$m = \frac{M}{\mu_0} \quad \text{and} \quad n = \frac{1}{2}(N + m).$$

Except when M is near its extreme possible values $\pm N\mu_0$ [where $P''(M)$ is negligibly small], the probability $P''(M)$ does not vary appreciably in going from one possible value of M to an adjacent one; i.e., $|P''(M + 2\mu_0) - P''(M)| \ll P''(M)$. The envelope of the possible values of $P''(M)$ then forms a smooth curve, as indicated in Fig. 2.11. Thus it is possible to regard $\overline{P''(M)}$ as a smoothly varying function of the continuous variable M , although only the discrete values (71) of M are relevant.

Suppose that μ_0 is negligibly small compared to the smallest magnetic moment of interest in any macroscopic measurement. The fact that M can assume only discrete values separated by $2\mu_0$ is then unobservable within the precision of contemplated observations. Thus M can indeed be considered as a continuous variable. Furthermore, one can meaningfully talk of a range dM which is a "macroscopic infinity."

+ Observable

in this example is
the total magnetic
moment M

exists

$P''(M) = P'(M)$
Smooth

$du = dM$

tesimal," i.e., a quantity which is *macroscopically* very small although it is *microscopically* large. (In other words, dM is supposed to be negligibly small compared to the smallest magnetic moment of interest in a macroscopic discussion, although it is much larger than μ_0 .)[†] The following question is then of interest: What is the probability that the total magnetic moment of the system lies in a particular small range between M and $M + dM$? The magnitude of this probability depends obviously on the size of the range dM and must become vanishingly small as dM is made negligibly small. This probability is therefore expected to be simply proportional to dM so that it can be written in the form:

$$\left[\begin{array}{l} \text{Probability that the total magnetic} \\ \text{moment lies between } M \text{ and } M + dM \end{array} \right] = \mathcal{P}(M) dM \quad (73)$$

where $\mathcal{P}(M)$ is independent of the magnitude of dM .[‡] The quantity $\mathcal{P}(M)$ is called a *probability density*; it yields an actual probability when it is multiplied by the infinitesimal range dM .

The probability (73) is readily expressed explicitly in terms of the probability $P''(M)$ that the total magnetic moment assumes the particular discrete value M . Since (71) shows that the possible values of M are separated by amounts $2\mu_0$ and since $dM \gg 2\mu_0$, the range between M and $M + dM$ contains $dM/(2\mu_0)$ possible values of M . All of these occur with nearly the same probability $P''(M)$ since the probability is very slowly varying over the small range dM . Hence the probability that the total moment lies in the range between M and $M + dM$ is simply obtained by summing $P''(M)$ over all values of M lying in this range, i.e., by multiplying the nearly constant value $P''(M)$ by $dM/(2\mu_0)$. This probability is thus properly proportional to dM and is given explicitly by

$$\mathcal{P}(M) dM = P''(M) \frac{dM}{2\mu_0} = \sum_{dM} P''(M) \quad (74)$$

[†] It is worth noting that many differentials used in physics are macroscopic infinitesimals. For example, in the study of electricity, one often talks of the charge Q on a body and of a charge increment dQ . This differential description is valid if dQ is understood to be much larger than the discrete electronic charge e , although it is supposed to be negligibly small compared to Q itself.

[‡] Since the probability is some smooth function of dM , it should be expressible near any value of M as a Taylor series in powers of dM when dM is small. Thus it should be of the form

$$\text{Probability} = a_0 + a_1 dM + a_2 (dM)^2 + \dots$$

where the coefficients $a_0, a_1, \dots$ depend on M . Here $a_0 = 0$ since the probability must approach zero as dM is made very small; furthermore, terms involving higher powers of dM are negligibly small compared to the leading term which is proportional to dM . Hence one is led to the result (73).

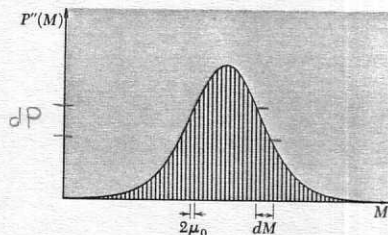


Fig. 2.11 The probability $P''(M)$ that the total magnetic moment of a spin system has a value M in a case where the number N of spins is large and the magnetic moment μ_0 of a spin is relatively small.

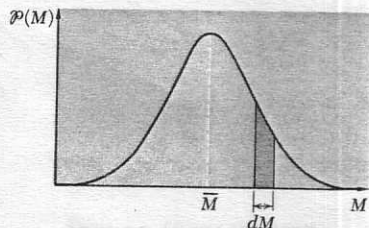


Fig. 2.12 The probability distribution of Fig. 2.11 is shown expressed in terms of the probability density $\mathcal{P}(M)$. Here $\mathcal{P}(M) dM$ [which is equal to the area under the curve in the small range between M and $M + dM$] is the probability that the total magnetic moment lies in the range between M and $M + dM$.

$$d\mathcal{P}(M, dM) = \mathcal{P}(M) dM$$

+ See, in (2.12) both M and dM lie on x.

.. The Probability between M and $M+dM$ depends on M and dM . Dependence on M is available in $\mathcal{P}(M)$ the density function

Observable

In practice, the actual calculation of $P'(M)$ may be laborious if M/μ_0 is large since the binomial distribution (14) then requires the evaluation of large factorials. These difficulties may, however, be circumvented by using the Gaussian approximation of Appendix A.1.

There are many problems where a variable of interest, call it u , is intrinsically continuous. For example, u might denote the angle between some vector in a plane and a fixed direction; this angle could then assume any value in the domain between 0 and 2π . In the general case u can assume any value in some domain $a_1 \leq u \leq a_2$. (This domain may be infinite in extent, i.e., $a_1 \rightarrow -\infty$, or $a_2 \rightarrow \infty$, or both.) Probability statements can be made about such a variable in a manner completely analogous to that discussed in the case of M . Thus one can focus attention on any infinitesimal range between u and $u + du$ and ask for the probability that the variable lies in this range. When du is sufficiently small, this probability must again be proportional to du so that it can be written in the form $P(u) du$ where the quantity $P(u)$ is a *probability density* independent of the size of du .

Probability considerations involving a continuous variable u can readily be reduced to the simpler situation where the possible values of the variables are discrete and thus countable. It is only necessary to subdivide the domain of possible values of u into arbitrarily small equal intervals of fixed size δu . Each such interval can then be labeled by some index r . The value of u in this interval can be denoted by u_r and the probability that u lies in this interval by P_r or $P(u_r)$. This procedure allows us to deal with a countable set of values of the variable u , each such value corresponding to one of the infinitesimal intervals $r = 1, 2, 3, \dots$. It also becomes apparent that relations involving probabilities of discrete variables remain equally valid for probabilities of continuous variables. For example, the simple properties (32) and (33) of mean values are also applicable if u is a continuous variable.

Note that the sums involved in calculating normalization conditions or mean values can be expressed as integrals if the variable is continuous. For example, the normalization condition asserts that the sum of the probabilities over all possible values of the variable must equal unity; in symbols

$$\sum_r P(u_r) = 1. \quad \text{Completed} \quad (75)$$

But, if the variable is continuous, one can first sum over all discrete intervals r for which u_r lies in the range between u and $u + du$; this gives the probability $P(u) du$ that the variable lies in this range.† One

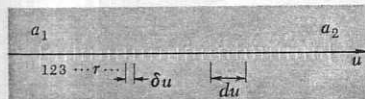


Fig. 2.13 Subdivision of the domain of a continuous variable u into a countable number of infinitesimal equal intervals of fixed size δu . Each such interval is labeled by an index r which can assume the values $1, 2, 3, 4, \dots$. The size of a macroscopic infinitesimal range du is also shown.

† Here the range du is understood to be large compared to the arbitrarily small interval δu (so that $du \gg \delta u$), but to be sufficiently small so that $P(u_r)$ does not vary appreciably within the range du .

can then complete the sum (75) by summing (i.e., integrating) over all such possible ranges du . Thus (75) is equivalent to

$$\int_{a_1}^{a_2} \mathcal{P}(u) du = 1 \quad (76)$$

which expresses the normalization condition in terms of the probability density $\mathcal{P}(u)$. Similarly, the general definition (31) of the mean value of a function $f(u)$ of discrete variables is given by

$$\overline{f(u)} \equiv \sum_r \mathcal{P}(u_r) f(u_r). \quad (77)$$

In a continuous description one can again sum first over all intervals r for which u_r lies in the range between u and $u + du$; this contributes to the sum an amount $\mathcal{P}(u) du f(u)$. One can then complete the sum by integrating over all possible ranges du . Hence (77) is equivalent to the relation†

$$\overline{f(u)} = \int_{a_1}^{a_2} \mathcal{P}(u) f(u) du. \quad (78)$$

Generalization to the case of several variables

is the Components of $E \cup$ other E 's

The generalization of the previous remarks to the case of more than one variable is immediate. Suppose, for example, that one is dealing with two continuous variables u and v . Then the joint probability that the variable u lies in the small range between u and $u + du$ and that the variable v lies in the small range between v and $v + dv$ is proportional to both du and dv and can be written in the form $\mathcal{P}(u, v) du dv$ where $\mathcal{P}(u, v)$ is a probability density independent of the size of du or dv . If desired, the situation can again be reduced to one of discrete probabilities by subdividing the variable u into very small fixed intervals δu , each of which can be labeled by some index r , and by subdividing the variable v into very small fixed intervals δv , each of which can be labeled by some other index s . In these terms the situation can then be described by specifying the probability P_{rs} that the variables have values lying in any given cell labeled by the pair of indices r and s .

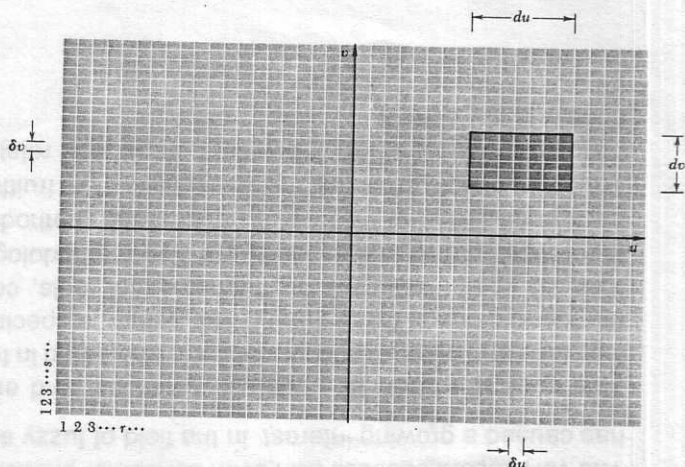


Fig. 2.14 Subdivision of the continuous variables u and v into small equal intervals of magnitude δu and δv labeled by r and s , respectively. The uv plane is thus subdivided into small cells, each labeled by the pair of indices r and s .

† Note that the probability density $\mathcal{P}(u)$ may become infinite for certain values of u . This does not lead to any difficulties as long as any integral $\int_{c_1}^{c_2} \mathcal{P}(u) du$ (which gives the probability that the value of u lies in an arbitrary range between c_1 and c_2) remains finite.

Summary of Definitions

statistical ensemble An assembly of a very large number of mutually noninteracting systems, each of which satisfies the same conditions as those known to be satisfied by a particular system under consideration.

time-independent ensemble An ensemble in which the number of systems exhibiting any particular property is the same at any time.

event The outcome of an experiment or result of an observation.

probability The probability P_r of occurrence of an event r in a system is defined with respect to a statistical ensemble of $\mathcal{N}$ such systems. If $\mathcal{N}_r$ systems in the ensemble exhibit the event r , then

$$P_r \equiv \frac{\mathcal{N}_r}{\mathcal{N}} \quad (\text{where } \mathcal{N} \rightarrow \infty).$$

statistical independence Two events are statistically independent if the occurrence of one event does not depend on the occurrence or nonoccurrence of the other event.

mean value (or ensemble average) The mean value of u is denoted by $\bar{u}$ and defined as

$$\bar{u} \equiv \sum_r P_r u_r$$

where the sum is over all possible values u_r of the variable u and where P_r denotes the probability of occurrence of the particular value u_r .

dispersion (or variance) The dispersion of u is defined as

$$(\Delta u)^2 \equiv \sum_r P_r (u_r - \bar{u})^2.$$

standard deviation The standard deviation of u is defined as

$$\Delta u \equiv [(\Delta u)^2]^{1/2}.$$

probability density The probability density $\mathcal{P}(u)$ is defined by the property that $\mathcal{P}(u) du$ yields the probability of finding the continuous variable u in the range between u and $u + du$.

Important Relations

Given N statistically independent events, each having probability p of occurring (and probability $q = 1 - p$ of not occurring).

Probability that n of these N events occur (binomial distribution):

$$P(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n}. \quad (\text{i})$$

Mean number of occurring events: $\bar{n} = Np. \quad (\text{ii})$

Standard deviation of n : $\Delta n = \sqrt{Npq}. \quad (\text{iii})$

Suggestions for Supplementary Reading

- W. Weaver, *Lady Luck* (Anchor Books, Doubleday & Company, Inc., Garden City, N.Y., 1963). An elementary introduction to probability concepts.
- F. Mosteller, R. E. K. Rourke, and G. B. Thomas, *Probability and Statistics* (Addison-Wesley Publishing Company, Reading, Mass., 1961).
- F. Reif, *Fundamentals of Statistical and Thermal Physics*, chap. 1 (McGraw-Hill Book Company, New York, 1965). The random walk problem discussed in this book is analogous to the problem of the ideal spin system, but is treated at greater length.
- H. D. Young, *Statistical Treatment of Experimental Data* (McGraw-Hill Book Company, New York, 1962). An elementary account of statistical methods, particularly as they are applied to problems of experimental measurement.
- W. Feller, *An Introduction to Probability Theory and its Applications*, 2d ed. (John Wiley and Sons, New York, 1959). This book on probability theory is more advanced than the preceding ones, but discusses many concrete examples.

Problems

2.1 An elementary dice problem

What is the probability of throwing a total of 6 points or less with 3 dice?

2.2 Random numbers

A number is chosen at random between 0 and 1. What is the probability that exactly 5 of its first 10 decimal places consist of digits less than 5?

2.3 Tossing of dice

Assume that each face of a die is equally likely to land uppermost. Consider a game which involves the tossing of 5 such dice. Find the probability that the number "6" appears uppermost

- (a) in exactly one die,
- (b) in at least one die,
- (c) in exactly two dice.

2.4 Probability of survival

In the macabre game of Russian roulette (*not* recommended by the author) one inserts a single cartridge into one of the 6 chambers of the cylinder of a revolver, leaving the other 5 chambers of the cylinder empty. One then spins the cylinder, aims at one's head, and pulls the trigger. What is the probability of still being alive after playing the game

- (a) Once?
- (b) Twice?
- (c) N times?
- (d) What is the probability of being shot when playing the game the N th time?

2.5 The random walk problem

A man starts out from a lamppost in the middle of a street, taking steps of equal length l . The probability is p that any one of his steps is to the right, and $q = 1 - p$ that it is to the left. The man is so drunk that his behavior at any step shows no traces of memory of what he did at preceding steps. His steps are thus statistically independent. Suppose that the man has taken N steps.

(a) What is the probability $P(n)$ that n of these steps are to the right and the remaining $n' = (N - n)$ steps are to the left?

(b) What is the probability $P'(m)$ that the displacement of the man from the lamppost is equal to ml , where $m = n - n'$ is an integer?

2.6 Probability of returning to starting point

In the preceding problem, suppose that $p = q$ so that each step is equally likely to be to the right or to the left. What is the probability that the man will again be at the lamppost after taking N steps

(a) If N is even?

(b) If N is odd?

2.7 One-dimensional diffusion of an atom

Consider a thin copper wire stretched out along the x axis. A few of the copper atoms, located near $x = 0$, are made radioactive by bombardment with fast particles. When the temperature of the wire is raised, the atoms become more mobile. Each atom can then jump to an adjacent lattice site, either to the site on its right (i.e., in the $+x$ direction) or the site on its left (i.e., in the $-x$ direction). The possible lattice sites are separated by a distance l . Assume that one has to wait a time τ before a given atom jumps to an adjacent lattice site. This time τ is a rapidly increasing function of the absolute temperature of the wire. The process whereby the atoms move about by virtue of successive jumps to adjacent sites is known as *diffusion*.

Suppose that the wire is rapidly raised to some high temperature at some time $t = 0$ and is thereafter maintained at this temperature.

(a) Let $P(x) dx$ denote the probability that a radioactive atom is found at a distance between x and $x + dx$ after a time t . [We assume that $t \gg \tau$ for all times t of physical interest, since τ is quite short when the temperature of the wire is high.] Make a rough sketch showing the behavior of $P(x)$ as a function of x in the following three cases:

(1) shortly after the time $t = 0$;

(2) after a moderately long time t has elapsed;

(3) after an extremely long time t has elapsed.

(b) What is the mean displacement $\bar{x}$ from the origin of a radioactive atom after a time t ?

(c) Find an explicit expression for the standard deviation Δx of the displacement of a radioactive atom after a time t .

2.8 Computation of the dispersion

Use the general properties of mean values to show that the dispersion of u can be calculated by the general relation

$$(\Delta u)^2 \equiv \overline{(u - \bar{u})^2} = \bar{u^2} - \bar{u}^2. \quad (i)$$

The last expression on the right provides a simple method for computing the dispersion.

Show also that (i) implies the general inequality

$$\bar{u^2} \geq \bar{u}^2. \quad (ii)$$

2.9 Mean values for a single spin

The magnetic moment of a spin $\frac{1}{2}$ is such that its component μ in the up direction has probability p of being equal to μ_0 , and probability $q = 1 - p$ of being equal to $-\mu_0$.

(a) Calculate $\bar{\mu}$ and $\bar{\mu^2}$.

(b) Use the expression (i) of Prob. 2.8 to calculate $\overline{(\Delta\mu)^2}$. Show that your result agrees with that given in Eq. (57) of the text.

2.10 The inequality $\bar{u^2} \geq \bar{u}^2$

Suppose that the variable u can assume the possible values u_r with respective probabilities P_r .

(a) By using the definitions of $\bar{u}$ and $\bar{u^2}$, and remembering the normalization requirement $\sum_r P_r = 1$, show that

$$\bar{u^2} - \bar{u}^2 = \frac{1}{2} \sum_r \sum_s P_r P_s (u_r - u_s)^2 \quad (i)$$

where each sum extends over all possible values of the variable u .

(b) Since no term in the sum (i) can ever be negative, show that

$$\bar{u^2} \geq \bar{u}^2 \quad (ii)$$

where the equals sign applies only in the case where only one value of u occurs with nonzero probability. The result (ii) agrees with that derived in Prob. 2.8.

2.11 The inequality $\overline{(u^n)^2} \leq \bar{u^{n+1}} \bar{u^{n-1}}$

The result (i) of the preceding problem suggests an immediate generalization. Thus consider the expression

$$\sum_r \sum_s P_r P_s u_r^m u_s^m (u_r - u_s)^2 \quad (i)$$

where m is any integer. When m is even, this expression can never be negative; when m is odd, it can never be negative if the possible values of u are all non-negative (or all nonpositive).

(a) By performing the indicated multiplications in (i), show the following results:

$$(\overline{u^n})^2 \leq \overline{u^{n+1}} \overline{u^{n-1}} \quad (\text{ii})$$

where $n \equiv m + 1$. If n is odd, this inequality is always valid; if n is even, it is valid if the possible values of u are all nonnegative (or all nonpositive). The equals sign in (ii) applies only in the case where a single possible value of u occurs with nonzero probability.

(b) Show that (ii) implies, as a special case, the inequality

$$\left(\overline{\frac{1}{u}}\right) \geq \frac{1}{\overline{u}}, \quad (\text{iii})$$

valid whenever the possible values of u are all positive (or all negative). The equals sign applies in the special case where only one value of u occurs with nonzero probability.

2.12 Optimum investment method

The following practical situation illustrates that different ways of averaging the same quantity can lead to significantly different results. Suppose that one desires to invest money by buying at the beginning of every month a certain number of shares of stock in a company. The cost c_r per share depends, of course, on the particular month r and varies from month to month in a fairly unpredictable fashion. Two alternative methods of regular investment suggest themselves: In method A, one buys the same number s of shares of stock every month; in method B, one buys stock for the same amount of money m every month. After N months, one will then have acquired a total number S of shares of stock and will have paid for them a total amount of money M . The best investment method is clearly that which yields the largest number of shares for the lowest amount of money, i.e., which yields the largest ratio S/M .

(a) Obtain an expression for the ratio S/M in the case of method A.

(b) Obtain an expression for the ratio S/M in the case of method B.

(c) Show that method B is the better investment method, no matter how the cost of shares fluctuates from month to month. [Suggestion: Exploit the inequality (iii) of the preceding problem.]

2.13 System of nuclei with spin 1

Consider a nucleus having spin 1 (i.e., spin angular momentum $\hbar$). Its component μ of magnetic moment along a given direction can then have three possible values, namely $+\mu_0$, 0, or $-\mu_0$. Suppose that the nucleus is not spherically symmetrical, but is ellipsoidal in shape. As a result it tends to be preferentially oriented so that its major axis is parallel to a given direction in the crystalline solid in which the nucleus is located. There is thus a probability p that $\mu = \mu_0$, and a probability p that $\mu = -\mu_0$; the probability that $\mu = 0$ is then equal to $1 - 2p$.

(a) Calculate $\bar{\mu}$ and $\overline{\mu^2}$.

(b) Calculate $\overline{(\Delta\mu)^2}$.

(c) Suppose that the solid under consideration contains N such nuclei which interact with each other to a negligible extent. Let M denote the total component of magnetic moment, along the specified direction, of all these nuclei. Calculate $\bar{M}$ and its standard deviation ΔM in terms of N , p , and μ_0 .

2.14 Direct calculation of $\bar{n}$ and $\overline{(\Delta n)^2}$

Consider an ideal system of N identical spins $\frac{1}{2}$. The number n of magnetic moments which point in the up direction can then be written in the form

$$n = u_1 + u_2 + \cdots + u_N \quad (i)$$

when $u_i = 1$ if the i th magnetic moment points up and $u_i = 0$ if it points down. Use the expression (i) and the fact that the spins are statistically independent to establish the following results.

(a) Show that $\bar{n} = N\bar{u}$.

(b) Show that $\overline{(\Delta n)^2} = N\overline{(\Delta u)^2}$.

(c) Suppose that a magnetic moment has probability p of pointing up and probability $q = 1 - p$ of pointing down. What are $\bar{u}$ and $\overline{(\Delta u)^2}$?

(d) Calculate $\bar{n}$ and $\overline{(\Delta n)^2}$ and show that your results agree with the relations (66) and (67) found in the text by a less direct method.

2.15 Density fluctuations in a gas

Consider an ideal gas of N molecules which is in equilibrium within a container of volume V_0 . Denote by n the number of molecules located within any subvolume V of this container. The probability p that a given molecule is located within this subvolume V is then given by $p = V/V_0$.

(a) What is the mean number $\bar{n}$ of molecules located within V ? Express your answer in terms of N , V_0 , and V .

(b) Find the standard deviation Δn in the number of molecules located within the subvolume V . Hence calculate $\Delta n/\bar{n}$, expressing your answer in terms of N , V_0 , and V .

(c) What does the answer to part (b) become when $V \ll V_0$?

(d) What value should the standard deviation Δn assume when $V \rightarrow V_0$? Does the answer to part (b) agree with this expectation?

2.16 Shot effect

Electrons of charge e are emitted at random from the hot filament of a vacuum tube. To good approximation the emission of any one electron does not affect the likelihood of emission of any other electron. Consider any very small time interval Δt . Then there is some probability p that an electron will be emitted from the filament during this time interval (and probability $q = 1 - p$ that an electron will not be emitted). Since Δt is very small, the probability p of emission during this time interval is very small (i.e., $p \ll 1$) and the probability that more than one electron will be emitted during the time Δt is negligible.

Consider any time t which is much larger than Δt . During this time there are then $N = t/\Delta t$ possible time intervals Δt during which an electron can be emitted. The total charge emitted in time t can then be written as

$$Q = q_1 + q_2 + q_3 + \cdots + q_N$$

where q_i denotes the charge which is emitted during the i th interval Δt ; thus $q_i = e$ if an electron is emitted and $q_i = 0$ if it is not.

(a) What is the mean charge $\bar{Q}$ emitted from the filament during the time t ?

(b) What is the dispersion $(\Delta Q)^2$ of the charge Q emitted from the filament during the time t ? Make use of the fact that $p \ll 1$ to simplify the answer to this question.

(c) The current I emitted during the time t is given by Q/t . Relate thus the dispersion $(\Delta I)^2$ of the current to the mean current $\bar{I}$, showing that

$$(\Delta I)^2 = \frac{e}{t} \bar{I}.$$

(d) The fact that the current measured during any interval t exhibits fluctuations (which become more pronounced the shorter the time interval, i.e., the fewer the total number of individual electrons involved in the emission process) is known as the *shot effect*. Calculate the standard deviation ΔI of the current if the mean current $\bar{I} = 1$ microampere and the measuring time is 1 second.

2.17 Calculation of a mean square value

A battery of total emf V is connected to a resistor R . As a result, an amount of power $P = V^2/R$ is dissipated in this resistor. The battery itself consists of N individual cells connected in series so that V is just equal to the sum of the emf's of all these cells. The battery is old, however, so that not all the cells are in perfect condition. Thus there is only a probability p that the emf of any individual cell has its normal value v ; and a probability $(1 - p)$ that the emf of any individual cell is zero because the cell has become internally shorted. The individual cells are statistically independent of each other. Under these conditions, calculate the mean power P dissipated in the resistor. Express the result in terms of N , v , p , and R .

2.18 Estimate of error of measurement

A man attempts to lay off a distance of 50 meters by placing a meter stick end to end 50 times in succession. This procedure is necessarily accompanied by some error. Thus the man cannot guarantee a distance of precisely one meter between the two chalk marks which he makes each time he places the meter stick on the ground. He knows, however, that the distance between the two marks is equally likely to lie anywhere between 99.8 and 100.2 cm, and that it certainly does not lie outside these limits. After repeating the operation 50 times, the man will indeed have laid off a mean distance of 50 meters. To estimate his total error, calculate the standard deviation of his measured distance.

2.19 Diffusion of a molecule in a gas

A molecule in a gas is free to move in three dimensions. Let s denote its displacement between successive collisions with other molecules. Displacements of the molecule between successive collisions are, to fair approximation, statistically independent. Furthermore, since there is no preferred direction in space, a molecule is as likely to move in a given direction as in the opposite direction. Thus its mean displacement $\bar{s} = 0$ (i.e., each component of this displacement vanishes on the average so that $\bar{s}_x = \bar{s}_y = \bar{s}_z = 0$).

The total displacement R of the molecule after N successive displacements can then be written as

$$R = s_1 + s_2 + s_3 + \cdots + s_N$$

where s_i denotes the i th displacement of the molecule. Use reasoning similar to that of Sec. 2.5 to answer the following questions:

- What is the mean displacement $\bar{R}$ of the molecule after N displacements?
- What is the standard deviation $\Delta R \equiv (\overline{(R - \bar{R})^2})^{1/2}$ of this displacement after N collisions? In particular, what is ΔR if the magnitude of each displacement s has the same length l ?

2.20 Displacement distribution of random oscillators

The displacement x of a classical simple harmonic oscillator as a function of the time t is given by

$$x = A \cos(\omega t + \varphi)$$

where ω is the angular frequency of the oscillator, A is its amplitude of oscillation, and φ is an arbitrary constant which can have any value in the range $0 \leq \varphi < 2\pi$. Suppose that one contemplates an ensemble of such oscillators all of which have the same frequency ω and amplitude A , but which have random phase relationships so that the probability that φ lies in the range between φ and $\varphi + d\varphi$ is given simply by $d\varphi/(2\pi)$. Find the probability $\mathcal{P}(x) dx$ that the displacement of an oscillator, at any given time t , is found to lie in the range between x and $x + dx$.

Chapter 3

Statistical Description of Systems of Particles

3.1	Specification of the State of a System	101
3.2	Statistical Ensemble	108
3.3	Statistical Postulates	111
3.4	Probability Calculations	116
3.5	Number of States Accessible to a Macroscopic System	118
3.6	Constraints, Equilibrium, and Irreversibility	124
3.7	Interaction between Systems	129
	Summary of Definitions	135
	Important Relations	136
	Suggestions for Supplementary Reading	136
	Problems	136

Chapter 3 Statistical Description of Systems of Particles

The preceding review of basic probability concepts has prepared us to turn the qualitative considerations of the first chapter into a systematic quantitative theory of systems consisting of very many particles. Our aim will be to combine statistical considerations with our knowledge of the laws of mechanics applicable to the particles constituting a macroscopic system. The resulting theory, therefore, is called *statistical mechanics*. The reasoning that leads to this theory is very simple and uses only the most primitive ideas of mechanics and probability. The beauty of the subject lies precisely in the fact that arguments of great simplicity and apparent innocence are capable of yielding results of impressive generality and predictive power.

Indeed, the arguments used to discuss a macroscopic system are completely analogous to those used to discuss the familiar experiment of tossing a set of coins. The essential ingredients for an analysis of this experiment are the following:

(i) Specification of the state of the system

We must have available a method for specifying any of the possible outcomes of an experiment involving the system. For example, the state of the set of coins after any toss can be described by specifying which particular face of each coin is uppermost.

(ii) Statistical ensemble

We have far too little information about the precise manner in which the coins are tossed to be able to use the laws of mechanics to make a unique prediction about the outcome of any particular experiment. Hence we resort to a statistical description of the situation. Instead of considering the particular set of coins of interest, we thus focus attention on an ensemble consisting of a very large number of similar sets of coins subjected to the same experiment. We can then ask for the probability of occurrence of any experimental outcome. This probability can be measured by observing the ensemble and determining the fraction of systems exhibiting the particular outcome. Our theoretical aim is the prediction of any such probability.

(iii) *Statistical postulates*

To make theoretical progress, it is necessary to introduce some postulates. In the case of ordinary coins of uniform density, there is nothing intrinsic in the laws of mechanics which implies that one face of a coin should land uppermost more frequently than the other. Hence we are led to introduce the *postulate* that "a priori" (i.e., based on our prior notions as yet unverified by actual observations) a coin has equal probability of landing on each of its two faces. This postulate is eminently reasonable and certainly does not contradict any of the laws of mechanics. The actual validity of the postulate can, however, only be decided by using it to make theoretical predictions and by checking that these predictions are confirmed by experimental observations. To the extent that such predictions are consistently verified, the validity of this postulate can be accepted with confidence.

(iv) *Probability calculations*

Once the basic postulate has been adopted, we can calculate the probability of occurrence of any particular outcome involving the set of coins under consideration. We can also compute various mean values of interest. Thus we can answer all questions that can meaningfully be asked in a statistical theory.

(C_i)_{3/4}

→ Reduction to Classical Mechanics
four Classical Thoughts C.M.

In studying systems consisting of a large number of particles, our considerations will be very similar to those used in formulating the preceding problem of a set of coins. The next four sections will make this analogy explicit.

3.1 *Specification of the State of a System*

The study of atomic particles has shown that any system of such particles is described by the laws of quantum mechanics. These laws, whose validity is supported by an overwhelming amount of experimental evidence, will thus form the conceptual basis for our entire discussion.

In a quantum-mechanical description the most precise possible measurement on a system always shows this system to be in some one of a set of discrete *quantum states* characteristic of the system. The microscopic state of a system can thus be described completely by specifying the particular quantum state in which the system is found.

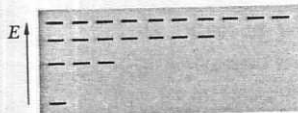


Fig. 3.1 Highly schematic diagram illustrating the first few energy levels of an arbitrary system. Each line denotes a possible quantum state of the system, while the vertical position of the line indicates the energy E of the system in this state. Note that there are many states which have the same energy.

r	σ	M	E
1	+1	μ_0	$-\mu_0 B$
2	-1	$-\mu_0$	$+\mu_0 B$

Table 3.1 Quantum states of a single spin $\frac{1}{2}$ having a magnetic moment μ_0 and located in a magnetic field B . Each state of the system can be labeled by an index r , or alternatively, by the quantum number σ . The magnetic moment (along the "up" direction specified by the field B) is denoted by M ; the total energy of the system is denoted by E .

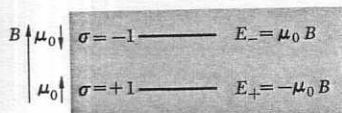


Fig. 3.2 Diagram showing the two energy levels of a spin $\frac{1}{2}$ having a magnetic moment μ_0 and located in a magnetic field B . The state where the magnetic moment points "up," so that its direction is parallel to B , is denoted by $\sigma = +1$ (or simply by $+$); that where it points "down" is denoted by $\sigma = -1$ (or simply by $-$).

Each quantum state of an isolated system is associated with a definite value of its energy and is called an *energy level*.[†] There may be several quantum states corresponding to the same energy of the system. (These quantum states are then said to be *degenerate*.) Every system has a lowest possible energy. There is usually only one possible quantum state of the system corresponding to this lowest energy; this state is said to be the *ground state* of the system.[‡] In addition there are, of course, many (indeed, ordinarily infinitely many) possible states with higher energies; these are called the *excited states* of the system. The preceding comments are completely general and applicable to any system, no matter how complex. They are best illustrated by some simple examples of great practical interest.

(i) Single spin

Consider a single particle, assumed to be fixed in position, which has spin $\frac{1}{2}$ and a magnetic moment of magnitude μ_0 . As already discussed in Sec. 1.3, this moment will be found to point either "up" or "down" (i.e., parallel or antiparallel) with respect to any specified direction. The system consisting of this single spin thus has only two quantum states which we shall label by a *quantum number* σ . We can then denote the state where the magnetic moment of the particle points up by $\sigma = +1$, and the state where it points down by $\sigma = -1$.

If the particle is in the presence of a magnetic field B , this field specifies the

direction of physical interest in the problem. The energy E of the system then is lower when the magnetic moment is aligned parallel to the field rather than antiparallel to it. The situation is analogous to that of a bar magnet located in an external magnetic field. Thus, when the magnetic moment points up (i.e., parallel to the field B), its magnetic energy is simply $-\mu_0 B$. Conversely, when the moment points down (i.e., antiparallel to the field B), its magnetic energy is simply $+\mu_0 B$. The two quantum states (or energy levels) of the system then correspond to different energies.

[†] The hydrogen atom is likely to be a familiar example of a system described in terms of discrete energy levels. Transitions of the atom between states of different energy give rise to the sharp spectral lines emitted by the atom. A description in terms of energy levels is, of course, equally applicable to any atom, molecule, or system consisting of many atoms.

[‡] In some cases there may be a relatively small number of quantum states of the same energy equal to the lowest possible energy of the system. The ground state of the system is then said to be *degenerate*.

(ii) Ideal system of N spins

Consider a system consisting of N particles, assumed to be fixed in position, where each particle has spin $\frac{1}{2}$ and a magnetic moment μ_0 . The system is located in an applied magnetic field $\mathbf{B}$. The interaction between the particles is assumed to be almost negligible.†

The magnetic moment of each particle can point either up or down with respect to the field $\mathbf{B}$. The orientation of the i th moment can thus be specified by the value of its quantum number σ_i so that $\sigma_i = +1$ when this moment points up and $\sigma_i = -1$ when it points down. A particular state of the whole system can then be specified by stating the orientation of

each of the N moments, i.e., by specifying the values assumed by the set of quantum numbers $\{\sigma_1, \sigma_2, \dots, \sigma_N\}$. Thus one can enumerate, and label by some index r , all the possible states of the whole system. This is done in Table 3.2 for the special case where $N = 4$. The total magnetic moment of the system is simply equal to the sum of the magnetic moments of the individual spins. Since the interaction between these spins is almost negligible, the total energy E of the system is also simply equal to the sum of the energies of the individual spins.

(iii) Particle in a one-dimensional box

Consider a single particle, of mass m , free to move in one dimension. The particle is supposed to be confined within a box of length L , so that the particle's position coordinate x must lie in the range $0 \leq x \leq L$. Within this box the particle is subject to no forces.

In a quantum-mechanical description, the particle has wave properties associated with it. The particle confined within the box and bouncing back and forth between its walls is thus represented by a wave function ψ in the form of a standing wave whose amplitude must vanish at the boundaries of the box (since ψ itself must vanish outside the box).‡ The wave function thus must be of the form

$$\psi(x) = A \sin Kx \quad (1)$$

(where A and K are constants) and must satisfy the boundary conditions

$$\psi(0) = 0 \quad \text{and} \quad \psi(L) = 0. \quad (2)$$

The expression (1) obviously satisfies the condition $\psi(0) = 0$. In order that it also satisfy the condition $\psi(L) = 0$, the constant K must be such that

$$KL = \pi n$$

or

$$K = \frac{\pi}{L} n, \quad (3)$$

where n can assume any of the integral values§

$$n = 1, 2, 3, 4, \dots \quad (4)$$

r	σ_1	σ_2	σ_3	σ_4	M	E
1	+	+	+	+	$4\mu_0$	$-4\mu_0 B$
2	+	+	+	-	$2\mu_0$	$-2\mu_0 B$
3	+	+	-	+	$2\mu_0$	$-2\mu_0 B$
4	+	-	+	+	$2\mu_0$	$-2\mu_0 B$
5	-	+	+	+	$2\mu_0$	$-2\mu_0 B$
6	+	+	-	-	0	0
7	+	-	+	-	0	0
8	+	-	-	+	0	0
9	-	+	+	-	0	0
10	-	+	-	+	0	0
11	-	-	+	+	0	0
12	+	-	-	-	$-2\mu_0$	$2\mu_0 B$
13	-	+	-	-	$-2\mu_0$	$2\mu_0 B$
14	-	-	+	-	$-2\mu_0$	$2\mu_0 B$
15	-	-	-	+	$-2\mu_0$	$2\mu_0 B$
16	-	-	-	-	$-4\mu_0$	$4\mu_0 B$

Table 3.2 Quantum states of an ideal system of 4 spins $\frac{1}{2}$, each having a magnetic moment μ_0 and located in a magnetic field $\mathbf{B}$. Each quantum state of the whole system is labeled by the index r , or equivalently, by the set of 4 numbers $\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$. For the sake of brevity, the symbol + indicates $\sigma = +1$ and the symbol - indicates $\sigma = -1$. The total magnetic moment (along the "up" direction specified by $\mathbf{B}$) is denoted by M ; the total energy of the system is denoted by E .

† This assumption implies that one can virtually neglect the magnetic field at the position of any one particle caused by the magnetic moments of the other particles.

‡ The physical interpretation of the wave function is that $|\psi(x)|^2 dx$ represents the probability that the particle is found in the range between x and $x + dx$.

§ The value $n = 0$ is irrelevant since it leads to $\psi = 0$, i.e., to no wave function (or no particle) existing within the box. Negative integral values of n lead to no distinct new wave functions since a change of sign of n , and thus of K , leads merely to a change of sign of ψ in (1) and leaves the probability $|\psi|^2 dx$ unchanged. Hence the positive integral values of n yield all the distinct possible wave functions of the form (1). Physically this means that only the magnitude $\hbar K$ of the particle's momentum is relevant, since this momentum is equally likely to be positive or negative as a result of successive reflections of the particle by the walls.

The constant K in (1) is the wave number associated with the particle; it is related to the wavelength λ (the so-called *de Broglie wavelength* associated with the particle) by the relation

$$K = \frac{2\pi}{\lambda}. \quad (5)$$

Hence (3) is equivalent to

$$L = n \frac{\lambda}{2}$$

and represents merely the familiar condition that standing waves are obtained when the length of the box is equal to some integral multiple of half-wavelengths.

The momentum p of the particle is related to K (or λ) by the famous de Broglie relation

$$p = \hbar K = \frac{h}{\lambda} \quad (6)$$

where $\hbar \equiv h/2\pi$ and h is Planck's constant. The energy E of the particle is simply its kinetic energy, since there is no potential energy due to external forces. Hence E can be expressed in terms of the velocity v or momentum $p = mv$ of the particle as

$$E = \frac{1}{2}mv^2 = \frac{1}{2}\frac{p^2}{m} = \frac{\hbar^2 K^2}{2m}. \quad (7)$$

The possible values (3) of K then yield the corresponding energies

$$E = \frac{\hbar^2}{2m} \left(\frac{\pi n}{L} \right)^2 = \frac{\pi^2 \hbar^2}{2m L^2} n^2. \quad (8)$$

Equivalently, we could have discussed the whole problem from a more mathematical point of view by starting from the fundamental Schrödinger equation for the wave function ψ . For a free particle in one dimension this equation is

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} = E\psi.$$

The functional form (1) satisfies this equation provided that the energy E is related to K by (7). The condition (2) that the wave function must vanish at the boundaries of the box leads again to (3) and hence to the expression (8) for the energy.

The possible quantum states of the particle in the box thus can be specified by the possible values (4) of the quantum number n . The corresponding discrete energies of these states (i.e., the corresponding energy levels of the particle) are then given by (8).

The relation (8) shows that the separation in energy between successive quantum states of the particle is very small if the length L of the box is of macroscopic size. The lowest possible energy of the particle, i.e., its ground-state energy, corresponds to the state $n = 1$. Note that this ground-state energy does not vanish.†

(iv) Particle in a three-dimensional box

The generalization of the preceding problem to the case of a particle free to move in three dimensions is straightforward. Suppose that the particle is confined within a box in the shape of a rectangular parallelepiped of edge lengths L_x , L_y , and L_z . The position coordinates x , y , z of the

particle then can be assumed to lie in the respective ranges

$$0 \leq x \leq L_x, \quad 0 \leq y \leq L_y, \quad 0 \leq z \leq L_z.$$

The particle has mass m and is subject to no forces within the box.

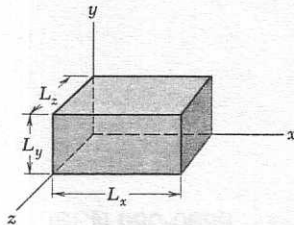


Fig. 3.3 A box in the shape of a rectangular parallelepiped with edge lengths L_x , L_y , and L_z .

† This conclusion is consistent with the Heisenberg uncertainty principle ($\Delta x \Delta p > \hbar$) according to which a particle confined within a linear dimension of length L (so that $\Delta x \sim L$) must have associated with it a minimum momentum p of the order of $p \sim \hbar/L$. Hence the lowest possible energy of the particle in the box is a kinetic energy of the order of $p^2/2m = \hbar^2/2mL^2$.

The wave function of the particle now represents a standing wave in three dimensions. Thus it is of the form

$$\psi = A(\sin K_x x)(\sin K_y y)(\sin K_z z) \quad (9)$$

where the constants K_x, K_y, K_z can be regarded as the three components of a vector $\mathbf{K}$, the *wave vector* of the particle. According to the de Broglie relation the momentum of the particle is then given by

$$\mathbf{p} = \hbar \mathbf{K} \quad (10)$$

so that the relationship between the magnitude of p and the magnitude of K (or the wavelength λ) is the same as in (6). The energy of the particle is then given by

$$\begin{aligned} E &= \frac{p^2}{2m} = \frac{\hbar^2 K^2}{2m} \\ &= \frac{\hbar^2}{2m} (K_x^2 + K_y^2 + K_z^2). \end{aligned} \quad (11)$$

Equivalently, it can be immediately verified that ψ in (9) is indeed a solution of the time-independent Schrödinger equation for a free particle in three dimensions,

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) = E\psi,$$

provided that E is related to $\mathbf{K}$ by (11).

The fact that ψ must vanish at the boundaries of the box imposes the conditions that

$\psi = 0$ at the planes

$$\left. \begin{aligned} x &= 0, & y &= 0, & z &= 0, \\ x &= L_x, & y &= L_y, & z &= L_z. \end{aligned} \right\} \quad (12)$$

The expression (9) vanishes properly when $x = 0$, $y = 0$, or $z = 0$. To make it vanish for $x = L_x$, $y = L_y$, or $z = L_z$, the constants K_x, K_y, K_z must satisfy the respective conditions

$$K_x = \frac{\pi}{L_x} n_x, \quad K_y = \frac{\pi}{L_y} n_y, \quad K_z = \frac{\pi}{L_z} n_z, \quad (13)$$

where each of the numbers n_x, n_y , and n_z can assume any of the positive integral values

$$n_x, n_y, n_z = 1, 2, 3, 4, \dots \quad (14)$$

Any particular quantum state of the particle can then be designated by the values assumed by the set of quantum numbers $\{n_x, n_y, n_z\}$. Its corresponding energy is, by (11) and (13), equal to

$$E = \frac{\pi^2 \hbar^2}{2m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right). \quad (15)$$

(v) Ideal gas of N particles in a box

Consider a system consisting of N particles, each of mass m , confined within the box of the preceding example. The interaction between the particles is assumed to be almost negligible so that the particles constitute an ideal gas. The total energy E of the gas is then just equal to the sum of the energies of the individual particles, i.e.,

$$E = \epsilon_1 + \epsilon_2 + \dots + \epsilon_N \quad (16)$$

where ϵ_i denotes the energy of the i th particle. The state of each such particle

can be specified, as in the preceding example, by the values of its 3 quantum numbers n_{ix}, n_{iy}, n_{iz} ; its energy ϵ_i is then given by an expression analogous to (15). Each possible quantum state of the *entire* gas can thus be specified by the values assumed by the $3N$ quantum numbers $\{n_{1x}, n_{1y}, n_{1z}; n_{2x}, n_{2y}, n_{2z}; \dots$

$$\dots; n_{Nx}, n_{Ny}, n_{Nz}\}.$$

Its corresponding energy is given by (16), where each term in the sum has the form (15).

The preceding examples are typical of the quantum-mechanical description and serve to illustrate the general comments made at the beginning of this section. We can summarize our remarks as follows:

1. Each possible quantum state of a system can be specified by some set of f quantum numbers. This number f , called the number of *degrees of freedom* of the system, is equal to the number of independent coordinates (including spin coordinates) needed to describe the system.[†]
2. Any one quantum state of the system can be specified by the particular values assumed by all of its quantum numbers. For simplicity, each such state can be labeled by some index r so that the possible quantum states can be listed and enumerated in some convenient order: $r = 1, 2, 3, 4, \dots$. Our question about the most detailed quantum-mechanical description of a system is then answered by the following statement:

The microscopic state of a system can be described by specifying the particular quantum state r in which the system is found.

3. A completely precise description of an isolated system of particles would have to take into account *all* the interactions between its particles and would determine the rigorously exact quantum states of the system. If the system is in any one of these exact states, it would then remain in this state forever. But no system is ever so completely isolated in practice that it does not interact at all with its environment; furthermore, it would be neither possible, nor indeed useful, to strive for precision so great that all interactions between particles are rigorously taken into account. Thus, the quantum states actually used to describe a system are in practice always its approximate quantum states, determined by taking into account all the important dynamical properties of its particles while neglecting small residual interactions.
4. A system initially known to be in one of its *approximate* quantum states does not remain in this state forever. In the course of time it will instead, under the influence of the small residual interactions, make transitions to its other quantum states (except to those to which it cannot pass without violating known restrictions imposed by the laws of mechanics).

[†]For instance, in the example of N particles without spin, the number of degrees of freedom $f = 3N$.

The hydrogen atom provides a familiar illustration of the preceding comments. The quantum states ordinarily used to describe this atom are those determined by taking into account only the Coulomb attraction between the nucleus and electron. The residual interaction of the atom with the surrounding electromagnetic field then causes transitions between these states. The result is the emission or absorption of electromagnetic radiation, giving rise to the observed spectral lines.

Illustrations of greater pertinence to us are provided by an isolated ideal spin system or an isolated ideal gas. If the particles in such a system would not interact with each other at all, the quantum states calculated in examples (ii) or (v) of this section would be exact, and hence no transitions would occur. But this situation does not correspond to reality. Indeed, we have been careful to stress that, even when a spin system or a gas is ideal, the interaction between its constituent particles is to be considered as *almost* negligible rather than as completely negligible. Thus small interactions exist in a spin system because each magnetic moment produces a small magnetic field at the positions of neighboring moments. Similarly, small interactions exist in the gas because forces exerted by one particle on another particle come into play whenever these particles approach each other sufficiently closely (and thus "collide" with each other). If these interactions are taken into account, the quantum states calculated in the examples (ii) and (v) become approximate quantum states. The effect of the interactions then is to cause occasional transitions between these states (the occurrence of such transitions being less frequent the smaller the magnitude of the interactions). Consider, for instance, the system consisting of 4 spins whose quantum states are listed in Table 3.2. Suppose that this system is originally known to be in the state $\{+ - + +\}$. As a result of the small interactions between the spins, there is a non-vanishing probability that the system will be found at some later time in any other state, such as $\{+ + - +\}$, to which it can pass without violating the condition of conservation of energy.

Our discussion of the specification of the state of a system has been given in terms of quantum ideas since the atoms and molecules constituting any system are known to be described by the laws of quantum mechanics. Under appropriate conditions it may sometimes be a convenient approximation to describe a system in terms of classical mechanics. The utility and validity of such approximations will be considered in Chap. 6.

Ideal interactions are almost negligible for each state parameter.

Isolated internal interactions can change a particle's state, and $E = \text{constant}$.

3.2 Statistical Ensemble

A precise knowledge of the particular microscopic state in which a system of particles is found at any one time would, in principle, allow us to use the laws of mechanics to calculate in the greatest possible detail all of the system's properties at any arbitrary time. Ordinarily, however, we do not have available such precise microscopic knowledge about a macroscopic system, nor are we interested in such an excessively detailed description. Hence we are led to discuss the system in terms of probability concepts. Instead of considering the single macroscopic system of interest, we then focus attention on an ensemble consisting of a very large number of systems of this kind, all satisfying the same conditions as those known to be satisfied by the particular system under consideration. With respect to this ensemble, we can then make various probability statements about the system.

A complete macroscopic description of a system of many particles defines the so-called *macroscopic state* or *macrostate* of the system. Since such a description is based entirely on the specification of quantities which can be readily ascertained by macroscopic measurements alone, it provides only very limited information about the particles in the system. This information is typically the following:

(Consider w.r.to G)

(i) Information about the external parameters of the system

There are certain parameters of the system which are macroscopically measurable and which affect the motion of the particles in the system. These parameters are called *external parameters* of the system. For example, the system may be known to be located in a given external magnetic field $\mathbf{B}$ or a given external electric field $\mathbf{E}$. Since the presence of such fields affects the motion of the particles in the system, $\mathbf{B}$ or $\mathbf{E}$ are external parameters of the system. Similarly, suppose that a gas is known to be contained within a box of dimensions L_x , L_y , and L_z . Then each molecule of the gas must move so as to remain confined within this box. The dimensions L_x , L_y , L_z are thus external parameters of the gas.

Since the external parameters affect the equations of motion of the particles in the system, they must also affect the energy levels of these particles. Thus the energy of each quantum state of a system is ordinarily a function of the external parameters of the system. For example, in the case of a spin system, Table 3.1 shows explicitly that the energies of the quantum states depend on the value of the external magnetic field B . Similarly, in the case of a particle in a box, the expression (15) shows explicitly that any quantum state specified by the

quantum numbers $\{n_x, n_y, n_z\}$ has an energy which depends on the dimensions L_x, L_y, L_z of the box.

A knowledge of all the external parameters of a system thus serves to determine the actual energies of its quantum states.

(ii) *Information about the initial preparation of the system*

- In view of the conservation laws of mechanics, the initial preparation of a system implies certain general restrictions on the subsequent motion of the particles in the system. For example, suppose that we are dealing with a system which is isolated so that it does not interact with any other system. Then the laws of mechanics require that the total energy of this system (i.e., the total kinetic and potential energy of all the particles in the system) remain constant. The system, when originally prepared for observation, must have some total energy which may be determined to within some finite precision, i.e., this energy may be known to lie within some small range between E and $E + \delta E$.
- † Then the conservation of energy implies that the total energy of the system must always be between E and $E + \delta E$. As a consequence of this restriction, the system can be found only in those of its quantum states which have an energy lying in this range.†

- We shall call the *accessible states* of a system those of its quantum states in which the system can be found without violating any conditions imposed by the information available about the system. A statistical ensemble selected in accordance with the information available about the system must, therefore, comprise systems all of which are in their accessible states. As pointed out previously, the specification of the macrostate of a system consisting of very many particles provides only very limited information about the system. If the system is known to be in a given *macrostate*, the number of quantum states accessible to the system is thus ordinarily very large (since the number

† In some cases one might also want to take into account other restrictions, such as those imposed by the conservation of total momentum. Although this could be done, it is ordinarily not of interest for the following reason: In the case of most laboratory experiments, it can be imagined that the system under investigation is enclosed in some container fastened to the floor of the laboratory and thus to the large mass of the earth. Any collision of the particles in the system with the container results then in a nearly negligible velocity change of the earth, although the latter can absorb any amount of momentum from the system while acquiring negligible energy. (The situation is similar to that of a ball being bounced off the earth.) Under these circumstances, there is no restriction on the possible momentum of the system, although its energy remains constant. Thus the system can be considered isolated with respect to energy transfer, while it is not isolated with respect to momentum transfer.

Parameters Space
 (1) external
 (2) macroscopically
 measurable

1. Fixed macroscopic measurement
 2. The ensemble reveals
 the accessible states
 a single system can attain

of particles in the system is very large). For instance, in the case of an isolated system merely known to have an energy between E and $E + \delta E$, all quantum states having energies in this range are accessible to the system.

It is conceptually simplest to discuss the case of a system which is *isolated* in the sense that it does not interact with any other systems so as to exchange energy with them.† Suppose that the macrostate of such an isolated system is specified by stating the values of its external parameters and the particular small range in which its energy is known to lie. This information determines then, respectively, the energies of the various quantum states of the system and the subset of those quantum states which are actually accessible to the system.

The essential content of the preceding remarks can be illustrated most simply by some systems with very few particles.

Example (i)

Consider a system of four spins $\frac{1}{2}$ (each having a magnetic moment μ_0) located in an applied magnetic field B . The possible quantum states and associated energies of this system are listed in Table 3.2. Suppose that this system is isolated and known to have a total energy which

is equal to $-2\mu_0 B$. The system can then be found in any one of the following four states accessible to it:

$$\begin{array}{ll} \{+++-\}, & \{++-+\}, \\ \{+-++\}, & \{-+++ \}. \end{array}$$

Example (ii)

Consider a system A^* which consists of two subsystems A and A' which can interact to a small extent and thus exchange energy with each other. System A consists of three spins $\frac{1}{2}$, each having a magnetic moment μ_0 . System A' consists of two spins $\frac{1}{2}$, each having a magnetic moment $2\mu_0$. The system A^* is located in an applied magnetic field B . We shall denote by M the total magnetic moment of A along the direction of B , and by M' the total magnetic moment of A' in this direction. The interaction between the spins is assumed to be almost negligible. The total energy E^* of the entire system A^* is then given by

$$E^* = -(M + M')B.$$

The system A^* consists of 5 spins and thus has a total of $2^5 = 32$ possible quantum states. Each of these can be labeled by five quantum numbers, the three numbers $\sigma_1, \sigma_2, \sigma_3$ specifying the orientations of the three magnetic moments of A , and the two quantum numbers σ'_1, σ'_2 specifying the orientations of the two magnetic moments of A' . Suppose that the isolated system A^* is known to have a total energy E^* equal to $-3\mu_0 B$. Then A^* must be found in any one of the five accessible states, listed in Table 3.3, which are compatible with this total energy.

r	σ_1	σ_2	σ_3	σ'_1	σ'_2	M	M'
1	+	+	+	+	-	$3\mu_0$	0
2	+	+	+	-	+	$3\mu_0$	0
3	+	-	-	+	+	$-\mu_0$	$4\mu_0$
4	-	+	-	+	+	$-\mu_0$	$4\mu_0$
5	-	-	+	+	+	$-\mu_0$	$4\mu_0$

Table 3.3 Systematic enumeration of all the states, labeled by some index r , which are accessible to the system A^* when its total energy in a magnetic field B is equal to $-3\mu_0 B$. The system A^* consists of a subsystem A with three spins $\frac{1}{2}$, each having magnetic moment μ_0 , and a subsystem A' with two spins $\frac{1}{2}$, each having magnetic moment $2\mu_0$.

† Any system which is *not* isolated can then be treated as a part of a larger system which is isolated.

The task of giving a statistical description of a macroscopic system can now be formulated very precisely. In a statistical ensemble of such systems, every system is known to be in one of its accessible quantum states. We should then like to predict the probability that the system is found in any given one of these accessible states. In particular, various macroscopic parameters of the system (e.g., its total magnetic moment or the pressure which it exerts) have values which depend on the particular quantum state of the system. Knowing the probability of finding the system in any one of its accessible states, we should then be able to answer the following questions of physical interest: What is the probability that any parameter of the system assumes any specified value? What is the mean value of such a parameter? What is its standard deviation?

3.3 Statistical Postulates

In order to make theoretical predictions concerning various probabilities and mean values, we must introduce some statistical postulates.

Let us, therefore, consider the simple case of an *isolated* system (with given external parameters) whose energy is known to lie in a specified small range between E and $E + \delta E$. As already mentioned, this system can be found in any one of a large number of accessible states. Focusing attention on a statistical ensemble of such systems, what can we say about the *probability* of finding the system in any one of these accessible states?

To elucidate this question, we shall exploit some simple physical arguments similar to those used in Secs. 1.1 and 1.2. There we considered the example of an ideal gas and talked about the distribution of molecules in a box over their possible positions in space. Analogously, our more abstract general arguments will now deal with the distribution of systems in an ensemble over their accessible states. Our discussion will then readily lead to the formulation of general postulates suitable as a basis for a statistical theory.

Let us examine first the simple situation where the system under consideration is known at some time to have equal probability of being found in each one of its accessible states. In other words, we consider the case where the systems in the statistical ensemble are known at some time to be uniformly distributed over all their accessible states. What happens then as time goes on? A system in a given state will, of course, not remain there forever; as we pointed out at the end of Sec. 3.1, it will continually make transitions between the various states

Always determine the space
1. Observed (internal)
2. Containing (external)
3. Observing
In terms of:
(Energy, Information,
accessible states)

distribution

Time and
Transitions

reveler dependent

$$\begin{array}{r} 1511510 \\ + 111111 \\ \hline \end{array}$$

accessible to it. The situation is thus a dynamic one. But there is nothing intrinsic in the laws of mechanics which gives preference to any one of the accessible states of a system compared to any other one. Thus, contemplating the ensemble of systems as time goes on, we do not expect that the number of systems in some particular subset of the accessible states will become less while that in some other subset will become greater.† Indeed, the laws of mechanics can be applied to an ensemble of isolated systems to show explicitly that, if the systems are initially uniformly distributed over all their accessible states, then they will remain uniformly distributed over these states forever.‡ Such a uniform distribution remains, therefore, unchanged in time.

Example

To give a very simple illustration, we may return to Example (i) of Sec. 3.2. There we dealt with an isolated system of four spins having a total energy $-2\mu_0 B$. Suppose that this system is known at some time to have equal probability of being found in each one of its four accessible states

$$\begin{array}{ll} \{+++-\}, & \{+-+-\}, \\ \{+--+ \}, & \{-+++\}. \end{array}$$

In accord with our preceding arguments, there is nothing intrinsic in the laws of

mechanics which would give preferred status to one of these four states compared to any other one. Hence, we do not expect that the system would at some later time be more likely to be found in one particular state, say the state $\{+++-\}$, rather than in any other one of its four accessible states. The contemplated situation, therefore, does not change in time. Hence the system will continue to have equal probability of being found in each one of its four accessible states.

The preceding argument thus leads to the following conclusion about an ensemble of isolated systems: If the systems in such an ensemble are uniformly distributed over their accessible states, the ensemble is time-independent. In terms of probabilities, this statement can be phrased as follows: If an isolated system has equal probability of being found in each one of its accessible states, then the probability of finding the system in each one of its states is time-independent.

† This argument is merely a more general version of that used in Sec. 1.1 where we considered an ideal gas. When the molecules of the gas are initially distributed uniformly throughout a box, they are not expected to concentrate themselves preferentially in one part of the box as time goes on since there is nothing in the laws of mechanics which gives preference to one part of the box compared to any other.

‡ This result is a consequence of the so-called "Liouville theorem." Proofs of this theorem, which require a knowledge of advanced mechanics far beyond the level of this book, can be found in R. C. Tolman, *The Principles of Statistical Mechanics*, chaps. 3 and 9 (Oxford University Press, Oxford, 1938).

An isolated system, by definition, is said to be *in equilibrium* if the probability of finding the system in each one of its accessible states is independent of time. In this case the mean value of every measurable macroscopic parameter of the system is, of course, also independent of time.† In terms of this definition of equilibrium, the conclusion of the preceding paragraph can be summarized by the following statement:

If an isolated system is found with equal probability in each one of its accessible states, it is in equilibrium. (17)

Let us next examine the general case where the isolated system under consideration is known at some initial time to be in some subset of the states actually accessible to it. A statistical ensemble of such systems at this time would then contain many systems in some subset of their accessible states, and no systems at all in the remaining accessible states. What happens now as time goes on? As we mentioned before, there is nothing intrinsic in the laws of mechanics which would tend to favor some of the accessible states over others. By their definition, the accessible states are also such that the laws of mechanics impose no restrictions preventing the system from being found in any one of these states. It is, therefore, exceedingly unlikely that a system in the ensemble remains indefinitely in the subset of states in which it finds itself initially and avoids the other states which are equally accessible to it.‡ Indeed, by virtue of small interactions between its constituent particles, the system will in the course of time continually make transitions between all of its accessible states. As a result, each system in the ensemble will ultimately pass through essentially all the states in which it can possibly be found. The net effect of these continual transitions is analogous to that resulting from the repeated shuffling of a deck of cards. If the shuffling is kept up long enough, the cards get so mixed up that each one is equally likely to occupy any position in the deck irrespective of how the deck was arranged initially. Similarly, in the case of the ensemble of systems, one expects that the systems will

† Indeed, to ascertain experimentally that a system is in equilibrium, one verifies that the mean values of all its observable macroscopic parameters are independent of time. One then presumes that the probability of finding the system in any one state is independent of time so that the system is in equilibrium.

‡ The present discussion in terms of quantum states and transitions between them is again merely a generalization of the arguments applied in Sec 1.2 to the case of an ideal gas. When all the molecules of the gas are in the left half of a box, the resulting situation is very unlikely and the molecules become quickly distributed throughout the entire box.

1 w.r to the ensemble
at same t : constant
2 w.r to one system
in the ensemble at t

Free evolution in time (a)

Laws of mechanics $\nRightarrow$ accessible states

Cause (a) = does not favor (a. states)

effect (b) = is not favored by L. of Mech.

ultimately become uniformly (i.e., randomly) distributed among all their accessible states.† Once this situation has been attained, the distribution remains uniform in accordance with (17). This final situation thus corresponds to a time-independent equilibrium situation.

The conclusion of the preceding argument can then be summarized as follows:

If an isolated system is not found with equal probability in each one of its accessible states, it is *not* in equilibrium. It then tends to change in time until it attains ultimately the equilibrium situation where it is found with equal probability in each one of its accessible states.

George

(18)

Note that this conclusion is analogous to the statement (1.7) of Chap. 1. It merely formulates in more precise and general terms the tendency of an isolated system to approach its most random situation.

Example

The preceding comments can again be illustrated with our preceding example of the isolated system of four spins. Suppose that this system has been prepared in such a way that it is initially known to be in the state $\{++++\}$. The total energy of the system is then $-2\mu_0 B$ and remains constant at this value. There are, however, three other states

$\{++--\}$, $\{+-++\}$, $\{-+++ \}$

which also have this energy and are thus equally accessible to the system. Indeed, as a result of small interactions between the magnetic moments, proc-

esses occur in which one moment flips from the "up" direction to the "down" direction while another moment does the reverse (the total energy remaining, of course, unchanged). As a result of any such process, the system goes from some initial accessible state to some other accessible state. After many repeated transitions of this kind, the system will ultimately be found with equal probability in any one of its four accessible states

$\{++--\}$, $\{+-++\}$,

$\{+---\}$, $\{-+++ \}$.

- + We shall adopt the statements (17) and (18) as the fundamental postulates of our statistical theory. Both (17) and (18) can actually be deduced from the laws of mechanics, (17) rigorously and (18) with the aid of some assumptions. The statement (18) is of special importance since it leads, in particular, to the following assertion:

† With some assumptions inherent in a statistical description, this expectation can be deduced from the laws of mechanics as a consequence of the so-called "H-theorem." A simple proof of this theorem and further references can be found in F. Reif, *Fundamentals of Statistical and Thermal Physics*, Appendix A.12 (McGraw-Hill Book Company, New York, 1965).

Notice: the ensemble is a representative of the behavior of any single system in it, hence, at t=0,

$$P(\sigma_i) \neq P(\sigma_j)$$

For any $i, j \Rightarrow$ the implied is the single system through the ensemble to add.

* If $i = 1/2$ only

$$\sigma_i = \sigma_1 = \text{up} = +1$$

$$\sigma_j = \sigma_2 = \text{down} = -1$$

If an isolated system is in equilibrium, it is found with equal probability in each one of its accessible states. (19)

This statement is the converse of (17). The validity of (19) follows immediately from (18). Indeed, if the conclusion of (19) were false, (18) implies that the premise of (19) would be violated.

The statistical situation of greatest simplicity is clearly that which is time-independent, i.e., that involving an isolated system in equilibrium. In this case, the statement (19) makes an unambiguous assertion about the probability of finding the system in each one of its accessible states. The statement (19) is, therefore, the fundamental postulate upon which we can erect the entire theory of macroscopic systems in equilibrium. This fundamental postulate of equilibrium statistical mechanics is sometimes called the *postulate of equal a priori probabilities*. Note that this postulate is exceedingly simple. Indeed, it is completely analogous to the simple postulate (assigning equal probabilities to head or tail) used in discussing the experiment of tossing a coin. The ultimate validity of the postulate (19) can only be established, of course, by checking its predictions against experimental observations. Since a large body of calculations based on this postulate have yielded results in very good agreement with experiment, the validity of this postulate can be accepted with great confidence.

The theoretical problems become much more complex when we are dealing with a statistical situation which *does* change with time, i.e., with a system which is *not* in equilibrium. In this case the only general statement which we have made is embodied in (18). This postulate makes an assertion about the direction in which the system tends to change (i.e., this direction is such that the system tends to approach the equilibrium situation of uniform statistical distribution over all accessible states). The postulate provides no information, however, about the actual time required for the system to attain its ultimate equilibrium condition (the so-called *relaxation time*). This time might be shorter than a microsecond or longer than a century, depending on the detailed nature of the interactions between the particles of the system and on the resultant frequency with which transitions actually occur between the accessible states of this system. A quantitative description of a nonequilibrium situation may thus become quite difficult since it requires a detailed analysis of how the probability of finding a system in each of its states changes with time. On the other hand, a problem involving a system in equilibrium merely requires use of the simple postulate (19) of equal a priori probabilities.

The Fundamental Postulate
is unambiguous.

Equilibrium $\Leftrightarrow P(G_i) = P(G_j)$
 $\forall i \neq j$ and $\forall t$

(Thur. 3-9-1932 10:20 AM)

we need a fundamental postulate of
general predictive power to lighten
to our brain thought the scientific
investigation to the whole world.
 $G_2 = (\text{Jesus})$ "I am the Light of
the world"
What makes it fundamental?
Why, how --

System Space of Contexts
dependent

Remarks concerning the applicability of equilibrium arguments

It is worth pointing out that the idealized concept of equilibrium is, in practice, a relative one. The important comparison is always between the relaxation time τ_r (the characteristic time necessary for the system to attain equilibrium when it is not in equilibrium initially) and the time τ_e of experimental interest in a given discussion.

Imagine, for example, that when the piston in Fig. 3.4 is suddenly pulled to the right, it takes about 10^{-3} seconds before equilibrium is attained and the whole gas is uniformly distributed throughout the entire volume of the box. Thus $\tau_r \sim 10^{-3}$ sec. Suppose now that, instead of doing the preceding experiment, we proceed as in Fig. 3.5 by moving the piston to the right quite slowly so that it takes a time $\tau_e = 100$ sec to complete the task. Strictly speaking, the gas is not in equilibrium during this time since its volume is changing. But since $\tau_e \gg \tau_r$, the molecules at any instant have had sufficient time to become essentially uniformly distributed throughout the entire volume available to them at that instant. The situation in the gas would then remain effectively unchanged

if one imagined the piston to be really stopped at any instant. In practice, the gas can thus be considered to be in equilibrium at all times.

As an example of the opposite limit where $\tau_e \ll \tau_r$, consider a piece of iron which is rusting at a very slow rate; in a time $\tau_r = 100$ years it would be completely transformed into iron oxide. Strictly speaking, this again is not an equilibrium situation. But during a time τ_e of experimental interest, say $\tau_e = 2$ days, the situation would be effectively unchanged if one imagined rusting was prevented (e.g., by removing all the surrounding oxygen gas). The iron would then be in strict equilibrium. Thus, in practice, the piece of iron can again be discussed by equilibrium arguments.

It is, therefore, only in the case when $\tau_e \sim \tau_r$ (i.e., when the time of experimental interest is comparable to the time required to reach equilibrium) that the time-dependence of the system is of essential significance. The problem is then more difficult and cannot be reduced to a discussion of equilibrium, or near-equilibrium, situations.

3.4 Probability Calculations

- The fundamental postulate (19) of equal a priori probabilities permits the statistical calculation of all time-independent properties of any system in equilibrium. In principle, such a calculation is very simple.
- Consider thus an isolated system in equilibrium and denote by Ω the total number of states accessible to it. According to our postulate, the probability of finding the system in each one of its accessible states is then the same and hence simply equal to $1/\Omega$. (The probability of finding the system in a state which is not accessible to it is, of course, zero.) Suppose now that we are interested in some parameter y of the system; for example, y might be the magnetic moment of the system or the pressure exerted by it. When the system is in any particular state, the parameter y assumes correspondingly some definite value.

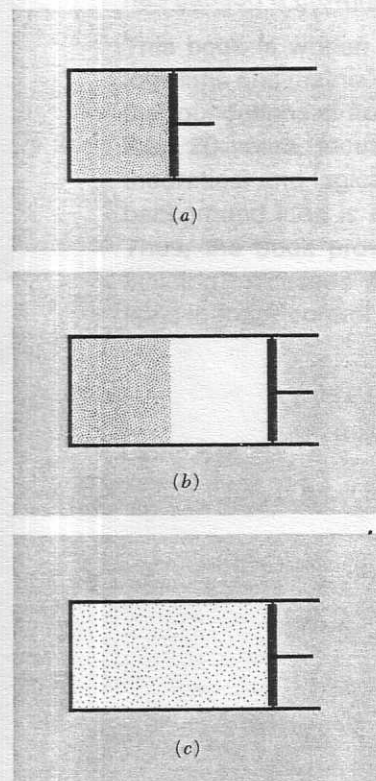


Fig. 3.4 Sudden expansion of a gas. (a) Initial situation. (b) Situation immediately after the piston has been moved. (c) Final situation.

† Let us enumerate the possible values which y can assume† and designate them by $y_1, y_2, \dots, y_n$. Among the Ω states accessible to the system, there will then be some Ω_i such states in which the parameter assumes the particular value y_i . The probability P_i that the parameter assumes the value y_i is then simply the probability that the system is found among the Ω_i states characterized by this value y_i . Thus P_i is obtained by summing $1/\Omega$ (the probability of finding the system in any single one of its accessible states) over the Ω_i states where y assumes the value y_i ; i.e., P_i is merely Ω_i times as large as the probability $1/\Omega$ of finding the system in each one of its accessible states. Hence†

$$P_i = \frac{\Omega_i}{\Omega}. \quad (20)$$

The mean value of the parameter y is then, by its definition, given by

$$\bar{y} \equiv \sum_{i=1}^n P_i y_i = \frac{1}{\Omega} \sum_{i=1}^n \Omega_i y_i \quad (21)$$

where the summation is over all possible values of y . The dispersion of y can be calculated in a similar way. All statistical calculations are thus basically just as simple as those involved in discussing the tossing of a set of coins.

Example (I)

Consider again the system of four spins whose states are listed in Table 3.2, and suppose that the total energy of this system is known to be $-2\mu_0 B$. If the system is in equilibrium, it is then equally likely to be in each one of the four states

$$\begin{aligned} \{++--\}, \quad \{+-+-\}, \\ \{-++-\}, \quad \{-+--\}. \end{aligned}$$

Focus attention on any one of these

spins, say the first. What is the probability P_+ that its magnetic moment points up? Since it points up in three of the four equally likely states accessible to the entire system, this probability is simply

$$P_+ = \frac{3}{4}.$$

What is the mean magnetic moment of this spin in the direction of the applied field B ? Since this moment is μ_0 in three

† If the possible values of the parameter are continuous rather than discrete, we can always proceed as in Sec. 2.6 and subdivide the domain of the possible values of y into very small intervals of fixed size δy . Since these intervals can be enumerated and labeled by some index i , we may then simply denote by y_i the value of the parameter when it lies anywhere in the i th such interval. The problem is then reduced to one where the possible values of y are discrete and countable.

† The result (20) is so simple because our fundamental postulate asserts that the system has equal probability of being found in each one of its accessible states. In an ensemble of $\mathcal{N}$ systems, the number $\mathcal{N}_i$ of systems with $y = y_i$ is therefore simply proportional to the number Ω_i of states accessible to a system with $y = y_i$. Thus $P_i \equiv \mathcal{N}_i/\mathcal{N} = \Omega_i/\Omega$.

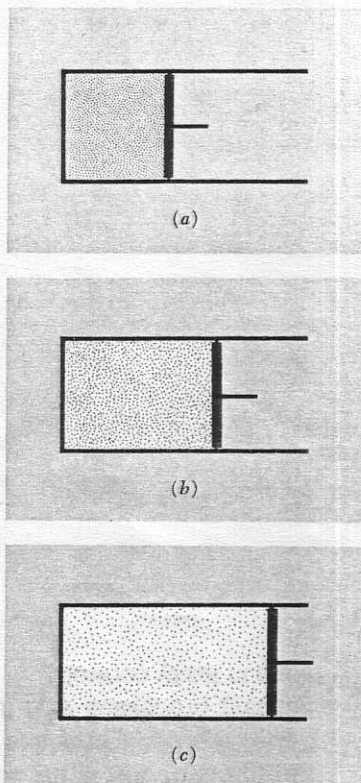


Fig. 3.5 Very slow (quasi-static) expansion of a gas. (a) Initial situation. (b) Intermediate situation. (c) Final situation.

1 initial describing DOF $\sim R$ coordinate
2 the reduced DOF ~ 1
 $\sim$ eigen states $r=1, 2, \dots$ representing a sequence of h^2
 $\in \mathbb{R}^2$

of the accessible states and $-\mu_0$ in one of the accessible states, this mean value is

$$\bar{M} = \frac{3\mu_0 + (-\mu_0)}{4} = \frac{1}{2}\mu_0.$$

† Note, incidentally, that one given spin in this system is not equally likely to have

a moment which points up or down; i.e., it is not found with equal probability in each one of its two possible states. This result does not, of course, contradict our fundamental statistical postulate since such a single spin is not isolated, but is part of a larger system where it is free to interact and exchange energy with other spins.

Example (ii)

Consider the system of spins discussed in Table 3.3. The total energy of this system is known to be $-3\mu_0 B$ and the system is assumed to be in equilibrium. The system is then equally likely to be found in each one of its 5 accessible states. Let us, for example, focus attention on the subsystem *A* consisting of 3 spins and denote by *M* its total magnetic moment in the direction of the field *B*. It is seen that *M* can assume the two

possible values $3\mu_0$ or $-\mu_0$. The probability $P(M)$ that it assumes any of these values can then be read off directly from the table; thus

$$P(3\mu_0) = \frac{2}{5},$$

and

$$P(-\mu_0) = \frac{3}{5}.$$

The mean value of *M* is then given by:

$$\bar{M} = \frac{(2)(3\mu_0) + (3)(-\mu_0)}{5} = \frac{3}{5}\mu_0.$$

† The preceding examples are extremely simple since they deal with systems consisting of very few particles. They do, however, illustrate the general procedure used in calculating probabilities and mean values for any system in equilibrium, no matter how complex. The only difference is that the enumeration of accessible states characterized by a particular value of a parameter may become a more difficult task in the case of a macroscopic system consisting of very many particles. The actual computation may thus be correspondingly more complicated.

3.5 *Number of States Accessible to a Macroscopic System*

The preceding four sections contain all the basic concepts needed for a quantitative statistical theory of macroscopic systems in equilibrium and for a qualitative discussion of their approach to equilibrium. We shall use the remainder of this chapter to become familiar with the significance of these concepts and to show how they provide a precise description of some of the qualitative notions introduced in Chap. 1. These preliminaries will prepare us to elaborate these concepts systematically throughout the rest of the book.

As we have seen, the properties of a system in equilibrium can be calculated by counting the number of states accessible to the system under various conditions. Although this counting problem may become difficult, it can ordinarily be circumvented in practice. As is usual in physics, progress may become easy if one strives to gain insight instead of trying to compute blindly. In particular, what is important in the present instance is to recognize some of the general properties exhibited by the number of states accessible to any system consisting of very many particles. Since a qualitative understanding of these properties and some approximate estimates will be quite adequate, fairly crude arguments will suffice to achieve our purpose.

Consider a macroscopic system with given external parameters so that its energy levels are determined. We shall denote the total energy of this system by E . To facilitate the counting of states, we shall group these states according to energy by subdividing the energy scale into equal small intervals of fixed magnitude δE . Thus δE is supposed to be very small on a macroscopic scale (i.e., very small compared to the total energy of the system and small compared to the contemplated precision of any macroscopic measurement of its energy). On the other hand, δE is supposed to be large on a microscopic scale (i.e., much larger than the energy of a single particle in the system and thus also much larger than the separation in energy between adjacent energy levels of the system). Any interval δE thus contains very many possible quantum states of the system. We shall introduce the notation

$$\Omega(E) \equiv \text{the number of states with energies lying in the interval between } E \text{ and } E + \delta E. \quad (22)$$

The number of states $\Omega(E)$ depends on the magnitude δE chosen as the subdivision interval in a given discussion. Since δE is macroscopically very small, $\Omega(E)$ must be simply proportional to δE , i.e., we can write†

$$\Omega(E) = \rho(E) \delta E \quad (23)$$

where $\rho(E)$ is independent of the size of δE . [The quantity $\rho(E)$ is called the *density of states* because it is equal to the number of states per unit energy range at the given energy E .] Since the interval δE

† The situation here is similar to that encountered in Sec. 2.6 when we discussed continuous probability distributions. The number of states $\Omega(E)$ must vanish when δE approaches zero and must be expressible as a Taylor series in powers of δE . When δE is sufficiently small, this series reduces to (23) since terms involving higher powers of δE are then negligible.

1st

2nd

(i)

(ii)

A Central Idea

Internal (observed) space

(i) macroscopic

(ii) microscopic

External (observing) space

contains very many states, $\Omega(E)$ changes only by a small fraction of itself as one goes from one energy interval to an adjacent one. Hence $\Omega(E)$ can be regarded as a smoothly varying function of the energy E . We shall be interested specifically in examining how sensitively $\Omega(E)$ depends on the energy E of a macroscopic system.

Note, incidentally, that it is possible to obtain $\Omega(E)$ if one knows the quantity

$$\Phi(E) \equiv \text{the total number of states having energies less than } E. \quad (24)$$

The number $\Omega(E)$ of states having an energy lying between E and $E + \delta E$ is then simply given by

$$\Omega(E) = \Phi(E + \delta E) - \Phi(E) = \frac{d\Phi}{dE} \delta E. \quad (25)$$

Before discussing the general properties of $\Omega(E)$ in the case of a macroscopic system, it will be instructive to illustrate how the number of states $\Omega(E)$ can be counted in the case of some very simple systems consisting of a single particle.

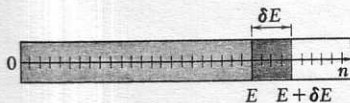


Fig. 3.6 The points on the line indicate the possible values $n = 1, 2, 3, 4, \dots$ of the quantum number n specifying the state of a single particle in one dimension. The values of n corresponding to the energy values E and $E + \delta E$ are indicated by the vertical lines. The region in light gray includes all the values of n for which the energy of the particle is less than E . The region in dark gray includes all the values of n for which the energy lies between E and $E + \delta E$.

Example (i) Single particle in a one-dimensional box

Consider a single particle of mass m free to move in one dimension in a box of length L . The possible energy levels of this system are then, by (8),

$$E = \frac{\hbar^2 \pi^2}{2m L^2} n^2 \quad (26)$$

where $n = 1, 2, 3, 4, \dots$. The coefficient of n^2 is very small if L is of macroscopic size. The quantum number n is thus very large for energies of ordinary interest.† By (26), the value of n for a given energy E is

$$n = \frac{L}{\pi \hbar} (2mE)^{1/2}. \quad (27)$$

Since successive quantum states correspond to values of n differing by unity, the total number $\Phi(E)$ of quantum states having an energy less than E , or a quantum number less than n , is then simply equal to $(n/1) = n$. Thus

$$\Phi(E) = n = \frac{L}{\pi \hbar} (2mE)^{1/2}. \quad (28)$$

Accordingly, (25) yields‡

$$\Omega(E) = \frac{L}{2\pi \hbar} (2m)^{1/2} E^{-1/2} \delta E. \quad (29)$$

† For example, if $L = 1$ cm and $m \sim 5 \times 10^{-23}$ gm, the mass of a nitrogen molecule given by (1.29), this coefficient is about 10^{-32} erg. But the average energy of such a molecule at room temperature is, by (1.28), of the order of 10^{-14} erg. Hence (26) yields for n a typical value of the order of 10^9 .

‡ Since n is very large, a change in n by unity produces only a negligible fractional change in n or E . The fact that n or E can assume only discrete values is then of negligible importance so that these variables can be treated as though they were continuous. In differentiating, it is merely necessary to consider any contemplated change in n to be always large compared to unity so that $dn > 1$, but to be very small in the sense that $dn \ll n$.

Example (ii) Single particle in a three-dimensional box

Consider a single particle of mass m free to move inside a three-dimensional box. For the sake of simplicity, we assume this box to be cubic and of edge length L . The possible energy levels of this system are then given by (15) with $L_x = L_y = L_z = L$; thus

$$E = \frac{\hbar^2}{2m} \frac{\pi^2}{L^2} (n_x^2 + n_y^2 + n_z^2) \quad (30)$$

where $n_x, n_y, n_z = 1, 2, 3, \dots$. In the "number space" defined by three perpendicular axes labeled by n_x, n_y, n_z , the possible values of these three quantum numbers lie thus geometrically at the centers of cubes of unit edge length, as indicated in Fig. 3.7. As in the preceding example, these quantum numbers are again ordinarily very large for a molecule in a macroscopic box. By (30) it follows that

$$n_x^2 + n_y^2 + n_z^2 = \left(\frac{L}{\pi\hbar}\right)^2 (2mE) \equiv R^2.$$

For a given value of E , the values of n_x, n_y, n_z which satisfy this equation lie on a sphere of radius R in Fig. 3.7. Here

$$R = \frac{L}{\pi\hbar} (2mE)^{1/2}.$$

The number $\Phi(E)$ of states with energy less than E is then equal to the number of unit cubes lying within this sphere and having positive values of n_x, n_y , and n_z ; i.e., it is simply equal to the volume of one octant of the sphere of radius R . Thus

$$\Phi(E) = \frac{1}{8} \left(\frac{4}{3} \pi R^3 \right) = \frac{\pi}{6} \left(\frac{L}{\pi\hbar} \right)^3 (2mE)^{3/2}. \quad (31)$$

By (25), the number of states with energy between E and $E + \delta E$ is then

$$\Omega(E) = \frac{V}{4\pi^2\hbar^3} (2m)^{3/2} E^{1/2} \delta E \quad (32)$$

where $V = L^3$ is the volume of the box.

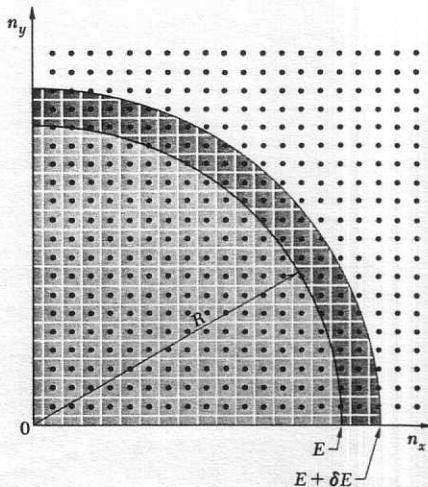


Fig. 3.7 The points indicate schematically in two dimensions the possible values of $n_x, n_y, n_z = 1, 2, 3, 4, \dots$ of the quantum numbers specifying the state of a single particle in three dimensions. (The n_z axis points out of the paper.) The values of n_x, n_y , and n_z corresponding to the energy values E and $E + \delta E$ lie on the indicated spherical surfaces. The region in light gray includes all values of n for which the energy of the particle is less than E . The region in dark gray includes all the values of n for which the energy lies between E and $E + \delta E$.

Let us now make a crude order-of-magnitude estimate to find roughly how the number of states $\Omega(E)$ or, equivalently $\Phi(E)$, depends on the energy E of a macroscopic system of particles. Any such system can be described by some set of f quantum numbers where f , the number of degrees of freedom of the system, is of the order of Avogadro's number. Associated with each quantum number there is a certain contribution ϵ to the total energy E of the system. We shall denote by $\varphi(E)$ the total number of possible values of this quantum number when its associated energy is less than ϵ . The number φ is then unity (or of the order of unity) when ϵ has its lowest possible value ϵ_0 , and must clearly increase as ϵ increases (although it may approach a constant value in an exceptional case).† Ordinarily we expect φ to increase roughly proportionately to the energy interval $(\epsilon - \epsilon_0)$. Hence we can write the approximate relation that

† This exceptional case arises if a system has only a finite total number of possible states and thus also an upper bound to its possible energy. (This can occur only if one ignores the kinetic energies of the particles in a system and considers only their spins.) Therefore, after at first increasing as a function of $(\epsilon - \epsilon_0)$, the quantity φ must in this case ultimately become constant.

$$\text{ordinarily, } \varphi(\epsilon) \propto (\epsilon - \epsilon_0)^\alpha \quad \text{where } \alpha \sim 1, \quad (33)$$

i.e., where α is some number of the order of unity.[†]

Consider now the entire system with f degrees of freedom. Its total energy E (the sum of the kinetic and potential energies of all the particles contained within the system) is the sum of the energies associated with all its degrees of freedom. Hence the value of the energy E (in excess of its lowest possible value E_0) should be roughly f times as large as the average energy ϵ per degree of freedom (in excess of its lowest possible value ϵ_0). Thus

$$E - E_0 \sim f(\epsilon - \epsilon_0). \quad (34)$$

Corresponding to a total energy of the system of E or less, there are then approximately $\varphi(\epsilon)$ possible values which can be assumed by the 1st of its quantum numbers, $\varphi(\epsilon)$ possible values which can be assumed by the 2nd of its quantum numbers, . . . , and $\varphi(\epsilon)$ possible values which can be assumed by the f th of its quantum numbers. The total number of possible combinations of these quantum numbers, i.e., the total number $\Phi(E)$ of states with total energy less than E , is then obtained by multiplying {the number of possible values of the 1st quantum number} by {the number of possible values of the 2nd quantum number}, then multiplying by {the number of possible values of the 3rd quantum number}, and finally multiplying by {the number of possible values of the f th quantum number}. Thus

$$\Phi(E) \sim [\varphi(\epsilon)]^f \quad (35)$$

where ϵ is related to E by (34). The number $\Omega(E)$ of states with energy between E and $E + \delta E$ is then given by (25). Thus

$$\Omega(E) = \frac{d\Phi}{dE} \delta E \sim f\varphi^{f-1} \frac{d\varphi}{dE} \delta E = \varphi^{f-1} \frac{d\varphi}{d\epsilon} \delta E \quad (36)$$

since $d\varphi/dE = f^{-1} (d\varphi/d\epsilon)$ by (34).

The preceding approximate considerations are quite sufficient to lead to some remarkable conclusions based on the fact that f is a very large number. Indeed, since we are dealing with a macroscopic system, f is of the order of Avogadro's number, i.e., $f \sim 10^{24}$. Numbers of this magnitude are so fantastically large that their peculiar properties are unlikely to be familiar from everyday experience.

As the energy E of the system increases, so does the energy ϵ per degree of freedom [see Eq. (34)]. Correspondingly, the number of

[†] For example, in the case of the quantum number describing the motion of a single particle in one dimension, $\varphi \propto \epsilon^{1/2}$ by virtue of (28). (The lowest possible energy ϵ_0 there is negligible compared to ϵ and thus is essentially zero.)

states $\varphi(\epsilon)$ per degree of freedom increases relatively slowly. But, since the exponents in (35) or (36) are of the order of f and thus extremely large, the number of states $\Phi(E)$, or $\Omega(E)$, of the system with f degrees of freedom increases at a fantastic rate. Hence we arrive at the following conclusion:

† The number of states $\Omega(E)$ accessible to any ordinary macroscopic system is an extremely rapidly increasing function of its energy E . (37)

growth ~ population dynamics in habitats

Indeed, combining (36) with (33) and (34), we obtain for the dependence of Ω on E the approximate relations

$$\Omega(E) \propto (\epsilon - \epsilon_0)^{\alpha f - 1} \propto \left(\frac{E - E_0}{f} \right)^{\alpha f - 1}.$$

Thus we can assert that,

for any ordinary system,
 $\Omega(E) \propto (E - E_0)^f$ approximately.† (38)

Here we have neglected 1 compared to f and have simply put $\alpha = 1$ since (38) is intended to indicate only very crudely the dependence of Ω on E . It matters little, for that purpose, whether the exponent in (38) is f , $\frac{1}{2}f$, or any other number of the order of f .

We can also make some statements about the magnitude of $\ln \Omega$. By (36) it follows that

$$\ln \Omega(E) = (f - 1) \ln \varphi + \ln \left(\frac{d\varphi}{d\epsilon} \delta E \right). \quad (39)$$

Let us now make the general observation that, if one deals with any large number such as f , its logarithm is always roughly of the order of unity and thus utterly negligible compared to the number itself. For example, if $f = 10^{24}$, $\ln f = 55$; thus $\ln f \ll f$. Consider then the terms on the right side of (39). The first term is of the order of f .†

† The qualifying phrase for any ordinary system is intended to exclude the exceptional case (mentioned in the footnote on page 121) where the kinetic energy of the particles in a system is ignored and where their spins have sufficiently high magnetic energy. (Focusing attention on the spins alone may be a justifiable approximation if the translational motion of the particles has only small effect on the orientations of their spins. The spin orientations and translational motion of the particles can then be treated separately.)

† This is always true unless the energy E of the system is very close to its ground state energy E_0 when $\varphi \sim 1$ for all degrees of freedom. Indeed, we know from our general comments at the beginning of Sec. 3.1 that, when a system approaches its ground state energy, the number of quantum states accessible to it is of order unity so that $\Omega \sim 1$.

The quantity $(d\varphi/d\epsilon) \delta E$ (where δE is any energy interval which is large compared to the spacing between the energy levels of the system) represents the number of possible values of a single quantum number in the interval δE and depends on the size of δE . But irrespective of the size chosen for δE , even the most conservative estimate would lead us to expect that this quantity $(d\varphi/d\epsilon) \delta E$ lies somewhere between 1 and 10^{100} . Its logarithm, however, then lies between 0 and 230 and is thus utterly negligible compared to f which is of the order of 10^{24} . Hence the second term on the right of (39) is utterly negligible compared to the first. We thus arrive at the following conclusions:

In the case of a macroscopic system, the number of states $\Omega(E)$ having an energy between E and $E + \delta E$ is to excellent approximation such that

$$\text{for } E \neq E_0, \quad \ln \Omega(E) \text{ is independent of } \delta E; \quad (40)$$

$$\ln \Omega(E) \sim f. \quad (41)$$

In words, (40) and (41) assert that, as long as the energy E is not very close to its ground state value, $\ln \Omega$ is independent of the magnitude of the chosen subdivision interval δE and is of the order of the number of degrees of freedom of the system.

3.6 Constraints, Equilibrium, and Irreversibility

Let us now summarize the general point of view which we have developed and which we shall apply repeatedly to discuss any situation involving macroscopic systems. The starting point for all our considerations is an isolated system.[†] Such a system is known to satisfy certain conditions which can be described on a macroscopic scale by specifying the value of some macroscopic parameter y of the system (or the values of several such parameters). These conditions act as constraints restricting the states in which the system can be found to those compatible with these conditions, i.e., to what we have called the *accessible states* of the system. The number Ω of these accessible states depends thus on the constraints to which the system is known to be subject so that $\Omega = \Omega(y)$ is some function of the specified macroscopic parameter of the system.

[†] Any system which is not isolated can always be treated as part of a larger system which is isolated.

Example

To keep in mind a specific illustration, consider the familiar system shown in Fig. 3.8a. This system consists of an ideal gas contained within the volume V_i constituting the left part of a box. The right part of the box is empty. Here the partition subdividing the box into the two

parts acts as a constraint which restricts the accessible states of the gas to those where all its molecules are located in the left part of the box. Hence the number Ω of accessible states of the gas depends on the volume V_i of this left part, i.e., $\Omega = \Omega(V_i)$.

The statistical description of the system involves probability statements about an *ensemble* of such systems all of which are subject to the same specified constraints. If the system is in equilibrium, it is found with equal probability in each one of its Ω accessible states, and conversely. If the system is not found with equal probability in each one of its Ω accessible states, then the statistical situation is not independent of time.[†] The system then tends to change until it ultimately attains the equilibrium situation where it is found with equal probability in each one of its Ω accessible states. The previous statements represent, of course, merely the content of our fundamental postulates (18) and (19).

Consider an isolated system initially in equilibrium under conditions where Ω_i states are accessible to it. The system is then found with equal probability in each one of these states. Suppose that some of the original constraints on the system are now removed. (For instance, in the example of Fig. 3.8, suppose that the partition is now removed.) Since the system is then subject to fewer restrictions than before, the number of states accessible to the system afterward certainly cannot be less than before; indeed, it will ordinarily be very much greater. Denoting by Ω_f this final number of states in the presence of the new constraints, we can thus write

$$\Omega_f \geq \Omega_i. \quad (42)$$

Immediately after the original constraints are removed, the probability of finding the system in any one of its states will be the same as before. Since the system had initially equal probability of being in each one of its Ω_i accessible states, it will thus, immediately after removal of the original constraints, still be found with equal probability in each one of these Ω_i states. Two cases can then arise:

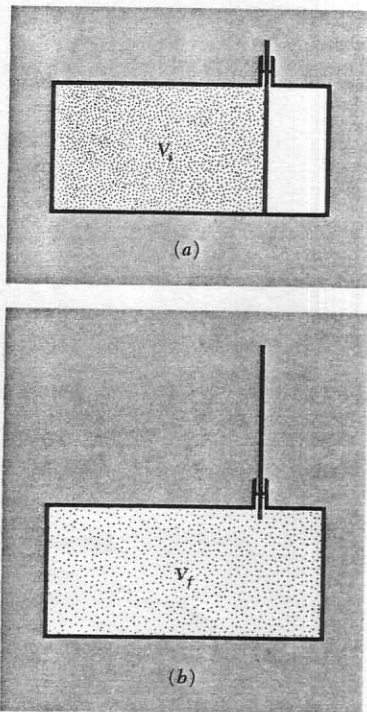


Fig. 3.8 An ideal gas confined within a box. (a) Initial situation where the gas is confined by a partition within the left volume V_i of the box. (b) Final situation, a long time after the partition is removed, when the gas is confined merely within the entire volume V_f of the box.

[†] In other words, in the ensemble of systems, the probability of finding the system in a given state changes with time for at least some of the states.

tion of the ensemble of systems cannot be restored merely by reimposing the former constraint while keeping the system isolated (i.e., while preventing the system from interacting with any other system with which it can exchange energy).

Of course, if we focus attention on a *single* system of the ensemble, its initial situation might be restored if we waited long enough for the occurrence of the right kind of spontaneous fluctuation. If this fluctuation is such that the system at a particular instant of time is found only among the Ω_i states originally accessible to it, then we could at that instant reimpose the original constraint and thus restore the system to its initial situation. But the probability of occurrence of such a fluctuation is ordinarily extremely small. Indeed, if we consider an ensemble of systems in equilibrium after the constraint has been removed, the probability P_i of encountering in the ensemble a system which is found among only Ω_i of its states is

$$P_i = \frac{\Omega_i}{\Omega_f} \quad (44)$$

since there are Ω_f states accessible to the system. The probability of encountering, among repeated observations of a *single* system, an observation corresponding to the desired fluctuation is then also given by (44). But in the usual case where $\Omega_f \gg \Omega_i$ [so that the final ensemble of systems differs very substantially from the initial one], (44) shows that spontaneous fluctuations allowing one to restore the initial situation of a single system occur exceedingly rarely.

We shall say that a process is *irreversible* if the initial situation of an *ensemble* of *isolated* systems having undergone this process cannot be restored by simply imposing a constraint. In accordance with this definition, the process in which an isolated system attains a new equilibrium situation after one of its constraints has been removed (thus changing the number of states accessible to it from Ω_i to Ω_f) is irreversible if $\Omega_f > \Omega_i$. Equivalently, our definition implies that a process is irreversible if an isolated system after this process has probability less than unity of being found in its original macrostate; indeed, in the usual case (where $\Omega_f \gg \Omega_i$) this probability is utterly negligible. Our present definition of irreversibility is thus merely a more precise formulation of that given in Sec. 1.2 in terms of the fluctuations in time of a single isolated system.

We can now also make more quantitative the notion of randomness introduced in Chap. 1. As a statistical measure of the degree of randomness of a system we can take the number of its accessible states actually occupied in an ensemble of such systems. The process of

attaining a new equilibrium situation after removing a constraint of an isolated system thus results in increasing the randomness of the system if $\Omega_f > \Omega_i$; the process is correspondingly irreversible.

Example

Once the gas in our previous example has attained its final equilibrium situation so that the molecules are essentially uniformly distributed throughout the entire box, the simple act of replacing the partition does not restore the original situation of an ensemble of such boxes. The molecules now located in the right part of the box still remain there. The process ensuing when the partition is removed is thus irreversible.

To examine possible fluctuations that might restore the original situation in a particular box of gas in the ensemble,

let us examine the probability P_i that this box is found with all its molecules on its left side after the final equilibrium situation has been attained. By virtue of (43) and (44), this probability is given by

$$P_i = \frac{\Omega_i}{\Omega_f} = \left(\frac{V_i}{V_f} \right)^N \quad (45)$$

and is thus fantastically small if $V_f > V_i$ and if N is at all large.[†] The removal of the partition thus illustrates a typical irreversible process in which the randomness of the gas is increased.

The general point of view of this section can be illuminated usefully by two further examples which are prototypes illustrating the interaction between macroscopic systems.

Example (i)

Consider an isolated system A^* which consists of two subsystems A and A' with fixed external parameters. (For instance, A and A' might be a piece of copper and a block of ice, respectively.) Suppose that A and A' are separated from each other in space so that they cannot exchange energy with each other. The system A^* is then subject to a constraint requiring that the energy E of A and the energy E' of A' must each separately remain constant. Accordingly, the states accessible to the total system A^* are only those consistent with the conditions that A has some constant specified energy E_i and that A' has some constant specified energy E'_i . If there are Ω_i^* such states accessible to A^* and if A^* is in equilibrium, then A^* is found with equal

probability in each one of these states.

Imagine that the systems A and A' are now placed in contact with each other so that they are free to exchange energy. Then the former constraint has been removed since the energies of A or A' need no longer be separately constant; only the sum $(E + E')$ of their energies, i.e., the total energy of the combined system A^* , must remain constant. As a result of removing the constraint, the number of states made accessible to A^* ordinarily becomes much greater, say, equal to Ω_f^* . Therefore, unless $\Omega_f^* = \Omega_i^*$, the system A^* is not in equilibrium immediately after A and A' are placed in contact with each other. The energies of the systems A and A' will, therefore, change (energy in the form of heat passing from one to the

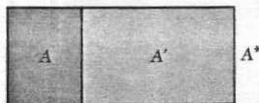


Fig. 3.9 Two systems A and A' with fixed external parameters and free to exchange energy. The combined system A^* , consisting of A and A' , is isolated.

[†] Note that in the special case where the box is originally divided by the partition into two halves, $V_f = 2V_i$ so that $P_i = 2^{-N}$. This is merely the result (1.1) obtained in Chap. 1 by primitive reasoning.

other) until A^* attains its final equilibrium situation where it is found with equal probability in each one of the Ω_f^* states now accessible to it.

Suppose that the systems A and A' are now separated again from each other so that they can no longer exchange energy. Although the former constraint

has thus been restored, the initial situation of A^* has not been restored (unless $\Omega_f^* = \Omega_i^*$). In particular, the mean energies of A and A' in the ensemble are now different from their initial values E_i and E'_i . The preceding process of heat transfer between the systems is thus irreversible.

Example (ii)

Consider an isolated system A^* consisting of two gases A and A' separated by a piston that is clamped in position. The piston then acts as a constraint requiring that the states accessible to A^* be only those consistent with the conditions that the molecules of gas A lie within some fixed volume V_i and that the molecules of gas A' lie within some fixed volume V'_i . If there are Ω_i^* such states accessible to A^* and if A^* is in equilibrium, then A^* is found with equal probability in each one of these states.

Imagine that the piston is now unclamped so that it is free to move. The individual volumes of the gases A and A' then no longer need be separately constant. Consequently, the number of states accessible to A^* becomes ordinarily much greater, say, equal to Ω_f^* . Unless $\Omega_f^* = \Omega_i^*$, the system A^* is thus not

in equilibrium after the piston is unclamped. Hence the piston will tend to move and the volumes of A and A' will correspondingly change until A^* attains its final equilibrium situation where it is found with equal probability in each one of the Ω_f^* states now accessible to it. As would be expected (and as will be shown explicitly in Chap. 6), the final volumes of the gases A and A' are then such that their mean pressures are equal, thus guaranteeing that the piston is indeed in mechanical equilibrium when it is in its final position.

The preceding process is again clearly irreversible if $\Omega_f^* > \Omega_i^*$. Leaving A^* isolated while simply clamping the piston again, so that it is not free to move, does not restore the initial volumes of the gases.

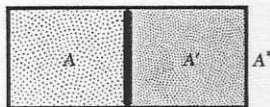


Fig. 3.10 Two gases A and A' separated by a movable piston. The combined system A^* , consisting of A and A' , is isolated.

3.7 Interaction between Systems

The preceding two examples illustrate specific instances where macroscopic systems interact with each other. Since the study of such interactions is of central importance,[†] we shall end this chapter by examining explicitly the distinct ways in which macroscopic systems can interact.

Consider two macroscopic systems A and A' which can interact and thus exchange energy with each other. The combined system A^* , consisting of A and A' , is then an isolated system whose total energy must remain constant. To describe the interaction between A and A' in statistical terms, we consider an ensemble composed of a very large

[†] Indeed, the whole discipline of *thermodynamics*, as the name indicates, deals with the macroscopic analysis of thermal and mechanical interactions and the macroscopic consequences that can be derived therefrom.

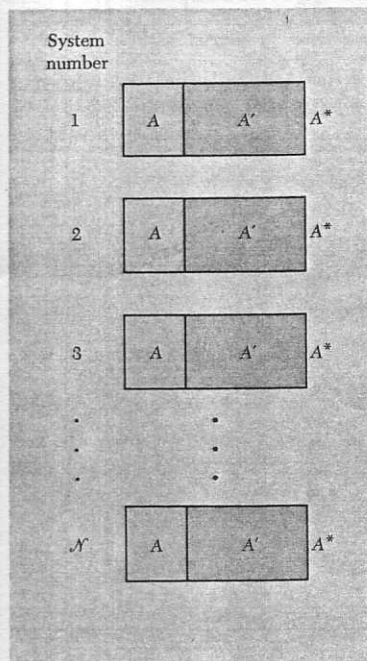


Fig. 3.11 An ensemble of systems A^* , each consisting of two systems A and A' capable of interacting with each other.

number of systems similar to A^* , each consisting of a pair of interacting systems A and A'. The process of interaction between A and A' does not ordinarily result in precisely the same energy transfer between A and A' for each such pair of systems in the ensemble. We can, however, meaningfully discuss the probability that an energy transfer of any specified magnitude results from the interaction process; or, more simply, we can merely ask for the *mean* energy transfer resulting from the interaction process. Let us then, in the ensemble of systems under consideration, denote the initial mean energies of A and A' before the interaction process by $\bar{E}_i$ and $\bar{E}'_i$, respectively; and let us denote the final mean energies of A and A' after the interaction process by $\bar{E}_f$ and $\bar{E}'_f$, respectively. Since the total energy of the isolated system A^* consisting of A and A' remains constant, it follows that

$$\bar{E}_f + \bar{E}'_f = \bar{E}_i + \bar{E}'_i; \quad (46)$$

i.e., the conservation of energy implies simply that

$$\Delta \bar{E} + \Delta \bar{E}' = 0 \quad (47)$$

where $\Delta \bar{E} \equiv \bar{E}_f - \bar{E}_i$ and $\Delta \bar{E}' \equiv \bar{E}'_f - \bar{E}'_i$ (48)

denote the changes of mean energy of each of the two systems A and A'.

We can now refine our discussion of Sec. 1.5 by examining systematically the various ways in which two macroscopic systems A and A' can interact. For this purpose, we shall inquire what happens to the external parameters of the systems during the interaction process.†

Thermal interaction

The interaction between the systems is particularly simple if all their external parameters are kept fixed so that their energy levels remain unchanged. We shall call such an interaction process *thermal interaction*. [Example (i) at the end of the preceding section provides a specific illustration.] The resulting increase (positive or negative) of the mean energy of a system is called the *heat absorbed* by the system and is conventionally denoted by the letter Q. Correspondingly, the decrease (positive or negative) of the mean energy of a system is called

† As mentioned at the beginning of Sec. 3.2, an external parameter of a system is a macroscopic parameter (such as the applied magnetic field B or the volume V) which affects the motion of the particles in the system and thus also the energy levels of this system. The energy E_r of each quantum state r of a system thus depends ordinarily on all the external parameters of the system.

the heat given off by the system and is given by $-Q$. Thus we can write

$$Q \equiv \Delta E \quad \text{and} \quad Q' \equiv \Delta E' \quad (49)$$

for the heat Q absorbed by system A and the heat Q' absorbed by system A' , respectively.† The conservation of energy (47) implies then that

$$Q + Q' = 0, \quad (50)$$

or

$$Q = -Q'.$$

This last relation asserts merely that the heat absorbed by A must be equal to the heat given off by A' . In accord with the definitions already introduced in connection with Eq. (1.15), the system absorbing a positive amount of heat is said to be the *colder* system, while the system absorbing a negative amount of heat (or giving off a positive amount of heat) is said to be the *warmer* or *hotter* system.

The characteristic feature of thermal interaction, where all external parameters are kept fixed, is that the energy levels of the systems remain unchanged while energy on the atomic scale is transferred from one system to the other. The mean energy of one of the systems increases then at the expense of the other *not* because the energies of its possible quantum states have changed, but because the system after the interaction is more likely to be found in those of its states which have higher energy.

Thermal insulation (or adiabatic isolation)

Thermal interaction between two systems can be prevented if they are properly separated from each other. Two systems are said to be *thermally insulated* or *adiabatically isolated* from each other if they cannot exchange energy as long as their external parameters are kept fixed.‡ Thermal insulation can be achieved by separating the systems sufficiently in space, or can be approximated by separating them from each other by a sufficiently thick partition made of a suitable material (such as asbestos or fiberglass). Such a partition is said to be *thermally insulating* or *adiabatic* if any two systems separated by it are found

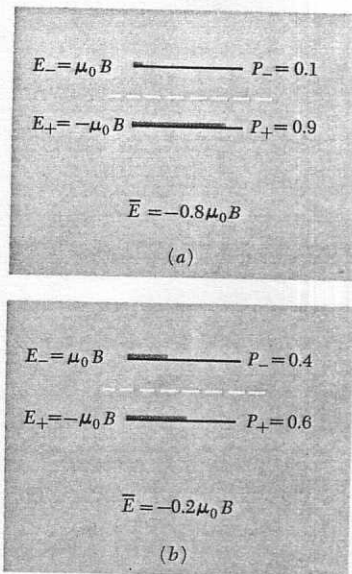


Fig. 3.12 The effect of thermal interaction on a very simple system A consisting of a single spin $\frac{1}{2}$ having a magnetic moment μ_0 and located in a magnetic field B . The diagram shows the two possible energy levels of A . These two quantum states are labeled by $+$ and $-$ and their corresponding energies are denoted by E_+ and E_- . The probability of finding A in a given state is denoted by P_+ and P_- , respectively, and its magnitude is indicated graphically by the length of the gray bar. The energy levels remain unchanged since the applied magnetic field (the one external parameter of this system) is supposed to be fixed. The initial equilibrium situation (a) is one where the spin is embedded in some solid. The solid, with its included spin, is then immersed in some liquid and one waits until the final equilibrium situation (b) is attained. In this process the spin system A absorbs heat from the system A' consisting of the solid and liquid. If the probabilities change as indicated in the diagram, the heat Q absorbed by A is equal to $Q = 0.6\mu_0 B$.

† Note that, since our present discussion has been more careful in its use of statistical concepts than that of Sec. 1.5, we have now defined heat in terms of the change of mean energy of a system.

‡ The word *adiabatic*, meaning "heat is not able to go through," comes from the Greek word *adiabatikos* (*a*, not + *dia*, through + *bainein*, to go). We shall always use the word in this sense, although it is sometimes also used in physics to denote a different concept.

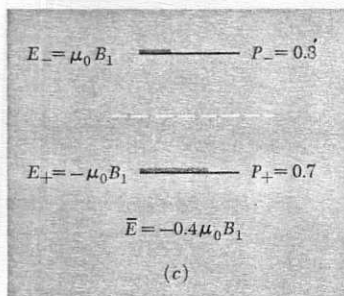
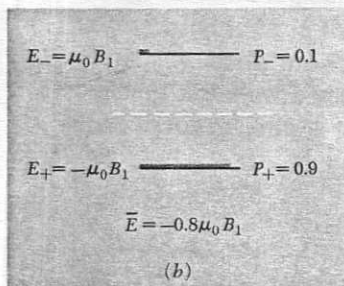
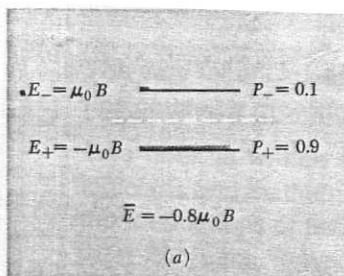


Fig. 3.13 The effect of adiabatic interaction on a very simple system A consisting of a single spin $\frac{1}{2}$ having a magnetic moment μ_0 and located in a magnetic field B . The initial situation and notation is the same as in Fig. 3.12, but the spin is then adiabatically isolated. Suppose that the magnetic field B is changed by means of an electromagnet. The amount of work done depends in general on just how the process is carried out. Panel (b) shows the final equilibrium situation if the magnetic field is changed from B to B_1 very slowly; the work W done on A is then $W = -0.8\mu_0(B_1 - B)$. Panel (c) shows a final equilibrium situation which might result if the magnetic field is changed from B to B_1 in some arbitrary way; the work done in the particular case shown here is $W = -0.4\mu_0 B_1 + 0.8\mu_0 B$.

to be thermally insulated from each other; i.e., if any such systems, initially in equilibrium, are found to remain in equilibrium as long as their external parameters are kept fixed.† A process which occurs while a system is thermally insulated from all other systems is called an *adiabatic process* of this system.

Adiabatic interaction

When two systems A and A' are thermally insulated from each other, they can still interact and thus exchange energy with each other, provided that at least some of their external parameters change in the process. We shall call such an interaction process *adiabatic interaction*. [Example (ii) at the end of Sec. 3.6 provides a specific illustration, if the piston is made of thermally insulating material.] The increase (positive or negative) of the mean energy of an adiabatically isolated system is called the *macroscopic work done on the system*‡ and will be denoted by W . Correspondingly, the decrease (positive or negative) of the mean energy of a system is called the *macroscopic work done by the system* and is given by $-W$. Hence we can write

$$W = \Delta \bar{E} \quad \text{and} \quad W' = \Delta \bar{E}' \quad (51)$$

for the work done on system A and the work done on system A' , respectively. If the total system $A + A'$ is isolated, the conservation of energy (47) implies then that

$$W + W' = 0 \quad (52)$$

or

$$W = -W'$$

This last relation asserts merely that the work done on one system must be equal to the work done by the other system.

Since adiabatic interaction involves changes in some of the external parameters of the systems, at least some of the energy levels of these systems are changed in the process. The mean energy of one of these interacting systems is thus ordinarily changed both because the energy of each of its states changes and because the probability that the system is found in any such state also changes.§

† If a partition is *not* thermally insulating, it is said to be *thermally conducting*.

‡ The macroscopic work, defined in terms of a mean energy difference, is a statistical quantity equal to the mean value of the work done on each system in the ensemble. Where no confusion is likely to arise, we shall henceforth simply use the term *work* to denote the macroscopic work thus defined.

§ It is worth noting a special case. If a system is in an exact quantum state whose energy depends on an external parameter, then the system will simply remain in this state and have its energy changed accordingly if the external parameter is changed *sufficiently slowly*.

General interaction

In the most general case, interacting systems are neither adiabatically isolated, nor are their external parameters kept fixed. It is then useful to write the total mean energy change of an interacting system, such as A , in the additive form

$$\Delta E = W + Q \quad (53)$$

where W represents the change of mean energy of A caused by changes of external parameters and Q represents the change of mean energy which is not caused by changes of external parameters. The decomposition (53) of ΔE into work W done on the system and heat Q absorbed by the system is meaningful when these contributions can be experimentally separated. Thus suppose that the system A interacts simultaneously with one system A'_1 which is separated from A by a thermally insulating partition, and another system A'_2 whose external parameters are kept fixed. Then the work W in (53) must simply be equal to the work done by (or decrease in mean energy of) the system A'_1 which is adiabatically isolated; similarly, the heat Q in (53) must simply be equal to the heat given off by (or decrease in mean energy of) the system A'_2 whose external parameters are kept fixed.

The general relation (53) is, for historical reasons, called the *first law of thermodynamics*. It explicitly recognizes work and heat as forms of energy transferred in different ways. Since both work and heat represent energies, these quantities are, of course, measured in units of energy, i.e., typically in terms of ergs or joules.†

Infinitesimal general interaction

An interaction process is particularly simple if it is infinitesimal in the sense that it takes a system from an initial macrostate to a final macrostate which differs from the initial macrostate only infinitesimally. The energy and external parameters of the system in the final macrostate differ then only very slightly from their values in the initial macrostate. Correspondingly, the infinitesimal increase in mean energy of the system can be written as the differential dE . In addition, we shall use the symbol dW , instead of W , to denote the infinitesimal amount of work done on the system in the process; this is merely a convenient notation designed to emphasize that the work is infinitesimally small. It is important to point out that dW does *not* designate any difference

† In the older literature of physics, and much of the present-day literature of chemistry, one still finds heat measured in terms of an old unit, the *calorie*, introduced in the eighteenth century before it was realized that heat is a form of energy. The calorie is now defined so that $1 \text{ calorie} \equiv 4.184 \text{ joules exactly}$.



Count Rumford (born Benjamin Thompson, 1753–1814). Born in Massachusetts, very much of an adventurer, he had strong sympathies during the American Revolution and left America to serve for a time as an aide to the Elector of Bavaria. His attention was attracted by the temperature rise produced by the boring of cannon barrels, an observation which led him to suggest that heat is merely a form of motion of particles in a body. Although this was an insight, it was too qualitative to have much impact on contemporary thinking which regarded heat as a substance ("phlogiston") conserved. (From a portrait in the Fogg Museum, painted in 1783 by T. Gainsborough, reproduced by courtesy of Harvard University.)

between works. Indeed, that would be a meaningless statement. The work done is a quantity referring to the interaction *process* itself; one *cannot* talk of the work in the system before and after the process, or of the difference between these. Similar comments apply to dQ which also denotes merely the infinitesimal amount of heat absorbed in a process, *not* any meaningless difference between heats. With this notation, the relation (53) for an infinitesimal process can be written in the form

$$dE = dW + dQ. \quad (54)$$

Remark

An infinitesimal process can be discussed particularly simply in statistical terms if it is carried out quasi-statically, i.e., so slowly that the system remains always very close to equilibrium. Let P_r denote the probability that the system A is in a state r having an energy E_r . The mean energy of this system is then, by definition,

$$\bar{E} = \sum_r P_r E_r \quad (55)$$

where the sum extends over all possible states r of the system. In an infinitesimal process the energies E_r change only by small amounts as a result of changes in the external parameters; furthermore, if the process is carried out very slowly, the probabilities P_r also change at most by small amounts. Hence the mean energy change in the process can be written as the differential

$$dE = \sum_r (P_r dE_r + E_r dP_r). \quad (56)$$

The heat absorbed corresponds to the increase of mean energy resulting when the external parameters are kept fixed, i.e., when the energy levels E_r are kept fixed so that $dE_r = 0$. Thus one can write

$$dQ = \sum_r E_r dP_r. \quad (57)$$

Hence the work done on the system is given by

$$dW = dE - dQ = \sum_r P_r dE_r. \quad (58)$$

This infinitesimal work is thus merely the mean energy change resulting from the shift of the energy levels produced by the infinitesimal change of external parameters, the probabilities P_r retaining their initial values appropriate to the equilibrium situation.

Summary of Definitions

(Some of these definitions are more precise than the definitions already encountered in previous chapters.)

microstate (or simply state) A particular quantum state of a system. It corresponds to the most detailed possible specification of a system described by quantum mechanics.

macrostate (or macroscopic state) A complete specification of a system in terms of macroscopically measurable parameters.

accessible state Any microstate in which a system can be found without contradicting the macroscopic information available about the system.

number of degrees of freedom The number of distinct quantum numbers necessary to describe completely the microstate of a system. It is equal to the number of independent coordinates (including spin coordinates) of all the particles in the system.

external parameter A macroscopically measurable parameter whose value affects the motion of the particles in a system and thus the energies of the possible quantum states of the system.

isolated system A system which does not interact with any other system so as to exchange energy with it.

total energy of a system The sum of the kinetic and potential energies of all the particles in the system.

internal energy of a system The total energy of a system measured in the frame of reference where its center of mass is at rest.

equilibrium An isolated system is said to be in equilibrium if the probability of finding the system in any one of its accessible states is independent of time. (The mean values of all macroscopic parameters of the system are then independent of time.)

constraint A macroscopic condition to which a system is known to be subject.

irreversible process A process which is such that the initial situation of an *ensemble* of isolated systems subjected to this process *cannot* be restored by simply imposing a constraint.

reversible process A process which is such that the initial situation of an *ensemble* of isolated systems subjected to this process *can* be restored by simply imposing a constraint.

thermal interaction Interaction in which the external parameters (and thus also the energy levels) of the interacting systems remain unchanged.

adiabatic isolation (or thermal insulation) A system is said to be adiabatically isolated (or thermally insulated) if it cannot interact thermally with any other system.

adiabatic interaction Interaction in which the interacting systems are adiabatically isolated. In this case the interaction process involves changes in some of the external parameters of the systems.

heat absorbed by a system Increase in mean energy of a system whose external parameters are kept fixed.

work done on a system Increase in mean energy of a system which is adiabatically isolated.

cold A comparative term applied to the system which absorbs positive heat as a result of thermal interaction with another system.

warm (or hot) A comparative term applied to the system which gives off positive heat as a result of thermal interaction with another system.



Fig. 3.15 Julius Robert Mayer (1814–1878). A German physician, Mayer suggested in 1842 the equivalence and conservation of all forms of energy, including heat. Although he made some quantitative estimates, his writing was too philosophical to be convincing and his work remained unrecognized for about twenty years. (From G. Holton and D. Roller, "Foundations of Modern Physical Science," Addison-Wesley Publishing Co., Inc., Cambridge, Mass., 1958. By permission of the publishers.)

Important Relations

Relation between mean energy, work, and heat:

$$\Delta E = W + Q \quad (i)$$

Suggestions for Supplementary Reading

Completely macroscopic discussions of heat, work, and energy:

M. W. Zemansky, *Heat and Thermodynamics*, 4th ed., secs. 3.1–3.5, 4.1–4.6 (McGraw-Hill Book Company, New York, 1957).

H. B. Callen, *Thermodynamics*, secs. 1.1–1.7 (John Wiley & Sons, Inc., New York, 1960).

Historical and biographical accounts:

C. Holton and D. Roller, *Foundations of Modern Physical Science* (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1958). Chapters 19 and 20 contain a historical account of the development of the ideas leading to the recognition of heat as a form of energy.

S. G. Brush, *Kinetic Theory*, vol. I (Pergamon Press, Oxford, 1965). The historical introduction by the author and the reprints of original papers by Mayer and Joule may be of particular interest.

S. B. Brown, *Count Rumford, Physicist Extraordinary* (Anchor Books, Doubleday & Company, Inc., Garden City, N.Y., 1962). A short biography of Count Rumford.

Problems

3.1 Simple example of thermal interaction*

Consider the system of spins described in Table 3.3. Suppose that, when the systems A and A' are initially separated from each other, measurements show the total magnetic moment of A to be $-3\mu_0$ and the total magnetic moment of A' to be $+4\mu_0$. The systems are now placed in thermal contact with each other and are allowed to exchange energy until the final equilibrium situation has been reached. Under these conditions calculate:

- (a) The probability $P(M)$ that the total magnetic moment of A assumes any one of its possible values M .
- (b) The mean value $\bar{M}$ of the total magnetic moment of A .
- (c) Suppose that the systems are now again separated so that they are no longer free to exchange energy with each other. What are the values of $P(M)$ and $\bar{M}$ of the system A after this separation?

3.2 One spin in thermal contact with a small spin system

Consider a system A consisting of a spin $\frac{1}{2}$ having magnetic moment μ_0 , and another system A' consisting of 3 spins $\frac{1}{2}$ each having magnetic moment μ_0 . Both systems are located in the same magnetic field B . The systems are placed in contact with each other so that they are free to exchange energy. Suppose that, when the moment of A points up (i.e., when A is in its $+$ state), two of the moments of A' point up and one of them points down. Count the total number of states accessible to the combined system $A + A'$ when the moment of A points up, and when it points down. Hence calculate the ratio P_-/P_+ , where P_- is the probability that the moment of A points down and P_+ is the probability that it points up. Assume that the total system $A + A'$ is isolated.

3.3 One spin in thermal contact with a large spin system

Generalize the preceding problem by considering the case where the system A' consists of some arbitrarily large number N of spins $\frac{1}{2}$, each having magnetic moment μ_0 . The system A consists again of a single spin $\frac{1}{2}$ with magnetic moment μ_0 . Both A and A' are located in the same magnetic field B and are placed in contact with each other so that they are free to exchange energy. When the moment of A points up, n of the moments of A' point up and the remaining $n' = N - n$ of the moments of A' point down.

(a) When the moment of A points up, find the number of states accessible to the combined system $A + A'$. This is, of course, just the number of ways in which the N spins of A' can be arranged so that n of them point up and n' of them point down.

(b) Suppose now that the moment of A points down. The total energy of the combined system $A + A'$ must, of course, remain unchanged. How many of the moments of A' now point up, and how many of them point down? Correspondingly, find the number of states accessible to the combined system $A + A'$.

(c) Calculate the ratio P_-/P_+ , where P_- is the probability that the moment of A points down and P_+ is the probability that it points up. Simplify your result by using the fact that $n \gg 1$ and $n' \gg 1$. Is the ratio P_-/P_+ larger or smaller than unity if $n > n'$?

3.4 Generalization of the preceding problem

Suppose that in the preceding problem the magnetic moment of A had the value $2\mu_0$. Calculate again the ratio P_-/P_+ of the probabilities that this moment points down or up.

3.5 Arbitrary system in thermal contact with a large spin system

The considerations of the preceding problems can readily be extended to treat the following general case. Consider any system A whatsoever, which might be a single atom or a macroscopic system. Suppose that this system A is placed in thermal contact with a system A' with which it is free to exchange energy. The system A' is supposed to be located in a magnetic field B and to consist of N spins $\frac{1}{2}$, each having a magnetic moment μ_0 . The number N is assumed to be very large compared to the number of degrees of freedom of the relatively much smaller system A . When the system A is in its lowest state of energy E_0 , we suppose that n of the moments of A' point up and the remaining $n' = N - n$ moments of A' point down. Here $n \gg 1$ and $n' \gg 1$, since all numbers are very large.

(a) When the system A is in its lowest possible state of energy E_0 , find the total number of states accessible to the combined system $A + A'$.

(b) Suppose now that the system A is in some other state, call it r , where it has an energy E_r higher than E_0 . In order to conserve the total energy of the combined system $A + A'$, there must then be $(n + \Delta n)$ moments of A' which point up, and $(n - \Delta n)$ moments of A' which point down. Express Δn in terms of the energy difference $(E_r - E_0)$. You can assume that $(E_r - E_0) \gg \mu_0 B$.

(c) When the system A is in the state r with energy E_r , find the total number of states accessible to the combined system $A + A'$.

(d) Let P_0 denote the probability that the system A is in the state of energy E_0 , and P_r the probability that it is in the state r of energy E_r . Find the ratio P_r/P_0 . Use the approximation that $\Delta n \ll n$ and $\Delta n \ll n'$.

(e) Using the results just derived, show that the probability P_r of finding the system A in any state r having an energy E_r is of the form

$$P_r = C e^{-\beta E_r}$$

where C is a constant of proportionality. Express β in terms of $\mu_0 B$ and the ratio n/n' .

(f) If $n > n'$, is β positive or negative? Suppose that the system A is such that its states, labeled by a quantum number r , are equally separated in energy by an amount b . (For example, A might be a simple harmonic oscillator.) Thus $\epsilon_r = a + br$, where $r = 0, 1, 2, 3, \dots$ and a is some constant. Compare the probability of finding the system A in any one of these states with the probability that it is found in its lowest state $r = 0$.

3.6 Pressure exerted by an ideal gas (quantum mechanical calculation)

Consider a single particle, of mass m , confined within a box of edge lengths L_x, L_y, L_z . Suppose that this particle is in a particular quantum state r specified by particular values of the three quantum numbers n_x, n_y, n_z . The energy E_r of this state is then given by (15).

When the particle is in the particular state r , it exerts on the right wall of the box (i.e., the wall $x = L_x$) some force F_r in the x direction. This wall must then

exert on the particle a force $-F_r$ (i.e., in the $-x$ direction). If the right wall of the box is slowly moved to the right by an amount dL_x , the work performed on the particle in this state is thus $-F_r dL_x$ and must be equal to the increase in energy dE_r of the particle in this state. Thus one has

$$dE_r = -F_r dL_x. \quad (i)$$

The force F_r exerted by a particle in the state r is thus related to the energy E_r of the particle in this state by

$$F_r = -\frac{\partial E_r}{\partial L_x}. \quad (ii)$$

Here we have written a partial derivative since the dimensions L_y and L_z are supposed to remain constant in deriving the expression (ii).

(a) Using (ii) and the expression (15) for the energy, calculate the force F exerted by the particle on the right wall when the particle is in a state specified by given values of n_x , n_y , and n_z .

(b) Suppose that the particle is not isolated, but is one of the many particles which constitute a gas confined within the container. The particle, being able to interact weakly with the other particles, can then be in any one of many possible states characterized by different values of n_x , n_y , and n_z . Express the mean force $\bar{F}$ exerted by the particle in terms of $\overline{n_x^2}$. For simplicity, assume that the box is cubic so that $L_x = L_y = L_z = L$; the symmetry of the situation then implies that $\overline{n_x^2} = \overline{n_y^2} = \overline{n_z^2}$. Use this result to relate $\bar{F}$ to the mean energy $\bar{E}$ of the particle.

(c) If there are N similar particles in the gas, the mean force exerted by all of them is simply $N\bar{F}$. Hence show that the mean pressure $\bar{p}$ of the gas (i.e., the mean force exerted by the gas per unit area of the wall) is simply given by

$$\bar{p} = \frac{2}{3} \frac{N}{V} \bar{E} \quad (iii)$$

where $\bar{E}$ is the mean energy of *one* particle in the gas.

(d) Note that the result (iii) agrees with that derived in (1.21) on the basis of approximate arguments using classical mechanics.

3.7 Typical number of states accessible to a gas molecule

The result (iii) of the preceding problem, or Eq. (1.21), permits one to estimate the mean energy of a gas molecule, such as nitrogen (N_2), at room temperature. Using the density and the known pressure exerted by such a gas, the mean energy $\bar{E}$ of such a molecule was found in (1.28) to be about 6×10^{-14} erg.

(a) Use (31) to calculate numerically the number of states $\Phi(\bar{E})$, with energy less than $\bar{E}$, accessible to such a molecule enclosed in a box having a volume of one liter (10^3 cm^3).

(b) Consider a small energy interval $\delta E = 10^{-24}$ erg, which is very much smaller than $\bar{E}$ itself. Calculate the number of states $\Omega(\bar{E})$ accessible to the molecule in the range between $\bar{E}$ and $\bar{E} + \delta E$.

(c) Show that the preceding number of states is very large, despite the small magnitude of the energy interval δE .

3.8 Number of states of an ideal gas

Consider an ideal gas consisting of N particles confined within a box with edge lengths L_x , L_y , and L_z . Here N is supposed to be of the order of Avogadro's number. By considering the energy contribution from each quantum number separately and using approximate arguments similar to those used in Sec. 3.5, show that the number of states $\Omega(E)$ in a given energy interval between E and $E + \delta E$ is given by

$$\Omega(E) = CV^N E^{(3/2)N} \delta E$$

where C is a constant of proportionality and $V = L_x L_y L_z$ is the volume of the box.

* 3.9 Number of states of a spin system

A system consists of N spins $\frac{1}{2}$, each having magnetic moment μ_0 , and is located in an applied magnetic field B . The system is of macroscopic size so that N is of the order of Avogadro's number. The energy of the system is then equal to

$$E = -(n - n')\mu_0 B$$

if n denotes the number of its magnetic moments which point up, and $n' = N - n$ the number which point down.

(a) Calculate for this spin system the number of states $\Omega(E)$ which lie in a small energy interval between E and $E + \delta E$. Here δE is understood to be large compared to individual spin energies, i.e., $\delta E \gg \mu_0 B$.

(b) Find an explicit expression for $\ln \Omega$ as a function of E . Since both n and n' are very large, apply the result $\ln n! \approx n \ln n - n$ derived in (M.10) to calculate both $n!$ and $n'!$. Show thus that, to excellent approximation,

$$\ln \Omega(E) = N \ln (2N) - \frac{1}{2}(N - E') \ln (N - E') - \frac{1}{2}(N + E') \ln (N + E')$$

where

$$E' \equiv \frac{E}{\mu_0 B}.$$

(c) Make a rough sketch showing the behavior of $\ln \Omega$ as a function of E . Note that $\Omega(E)$ does not always increase as a function of E . The reason is that a system of spins is anomalous in that it has not only a lowest possible energy $E = -N\mu_0 B$, but also a highest possible energy $E = N\mu_0 B$. On the other hand, in all ordinary systems where one does not ignore the kinetic energy of the particles (as we did in discussing the spins), there is no upper bound on the magnitude of the kinetic energy of the system.

Chapter 4

Thermal Interaction

4.1	Distribution of Energy between Macroscopic Systems	142
4.2	The Approach to Thermal Equilibrium	147
4.3	Temperature	149
4.4	Small Heat Transfer	155
4.5	System in Contact with a Heat Reservoir	157
4.6	Paramagnetism	163
4.7	Mean Energy of an Ideal Gas	166
4.8	Mean Pressure of an Ideal Gas	172
	Summary of Definitions	176
	Important Relations	177
	Suggestions for Supplementary Reading	177
	Problems	178

Chapter 4 Thermal Interaction

The preceding chapter has provided us with all the postulates and essential theoretical framework necessary for a quantitative discussion of macroscopic systems. We are, therefore, ready to explore the power of our concepts by applying them to some problems of major physical importance.

We shall begin by examining in detail the thermal interaction between systems. The analysis of this situation is particularly simple since the external parameters, and hence also energy levels, of the systems remain fixed. In addition, thermal interaction is one of the processes occurring most commonly in the world around us. The particular questions which we shall want to investigate are the following:

- 1. What conditions must be satisfied so that two systems in thermal interaction with each other are in equilibrium? What happens when these conditions are not satisfied? What probability statements can be made? We shall see that these questions can be answered quite simply to yield results of remarkable generality and great utility. Indeed, in this chapter we shall attain a good understanding of the concept of temperature and a precise definition of the "absolute temperature."
- 2. Furthermore, we shall obtain some eminently practical methods for calculating the properties of any macroscopic system in equilibrium on the basis of our knowledge about the atoms or molecules of which it is composed. Finally, we shall apply these methods to deduce explicitly the macroscopic properties of some specific systems.

4.1 Distribution of Energy between Macroscopic Systems

Consider two macroscopic systems A and A' . We shall denote the energies of these systems by E and E' , respectively. To facilitate the counting of states, we proceed as in Sec. 3.5 and imagine the energy scales to be subdivided into very small equal intervals of fixed size δE . (The magnitude of δE is, however, supposed to be large enough to contain many states.) We shall then denote by $\Omega(E)$ the number of states accessible to A when its energy lies between E and $E + \delta E$, and by $\Omega'(E')$ the number of states accessible to A' when its energy lies between E' and $E' + \delta E$. The counting problem can be simplified since we can, to excellent approximation, treat all energies as if they

Aim

+ In the Internal (Local) Space

$$\begin{array}{c} \text{(variable is)} \quad E, E' \\ \hline E + \delta E, E' + \delta E \\ \quad ? \qquad \quad ? \end{array}$$

could assume only discrete values separated by the small amount δE .

In particular, we can lump together all states of A which have an energy in the small interval between E and $E + \delta E$ and treat them as if they simply had an energy equal to E ; there are thus $\Omega(E)$ such states. Similarly, we can lump together all states of A' which have an energy in the small interval between E' and $E' + \delta E$ and treat them as if they simply had an energy equal to E' ; there are thus $\Omega'(E')$ such states. If we adopt this procedure, the statement that A has an energy E means physically that A has an energy anywhere between E and $E + \delta E$. Similarly, the statement that A' has an energy E' means physically that A' has an energy anywhere between E' and $E' + \delta E$.

(2) The systems A and A' have fixed external parameters, but are assumed to be free to exchange energy. (Any energy transfer between them is thus, by definition, in the form of heat.) Although the energy of each system separately is then not constant, the combined system $A^* \equiv A + A'$ is isolated so that its total energy E^* must remain unchanged. Hence†

$$E + E' = E^* = \text{constant}. \quad (1)$$

When the energy of A is known to be equal to E , the energy of A' is then determined to be

$$E' = E^* - E. \quad (2)$$

(3) Now consider the situation where A and A' are in equilibrium with each other, i.e., where the combined system A^* is in equilibrium. The energy of A can then assume many possible values. The question of interest is, however, the following: What is the probability $P(E)$ that the energy of A is equal to E (i.e., that it lies in the interval between E and $E + \delta E$) where E has any specified value? [The energy of A' , of course, then has a corresponding value E' given by (2).] The answer to this question is readily obtained by focusing attention on the combined isolated system A^* , since the basic postulate (3.19) asserts that such a system is equally likely to be found in each one of its accessible states. We have merely to pose the following question: Among the total number Ω_{tot}^* of states accessible to A^* , what is the number $\Omega^*(E)$ of states of A^* which are such that the subsystem A has an energy

† In our discussion E denotes the energy of A irrespective of A' , and E' denotes the energy of A' irrespective of A . The total energy E^* written as the simple sum (1) neglects, therefore, any interaction energy E_i which depends on both A and A' , i.e., any work necessary to bring the systems together. By definition, thermal interaction is supposed to be sufficiently weak so that E_i is negligible, i.e., so that $E_i \ll E$ and $E_i \ll E'$.

ie instantaneous state of the system as a whole. due to the (δE) transmitted difference in the whole system is transmitted to all constituents of the system.

How the constituents of A^* realize that they must raise each other energy so always (3) is satisfied? $\rightarrow$ to communicate E to excite, and the excitation is dependent on the source of excitation.

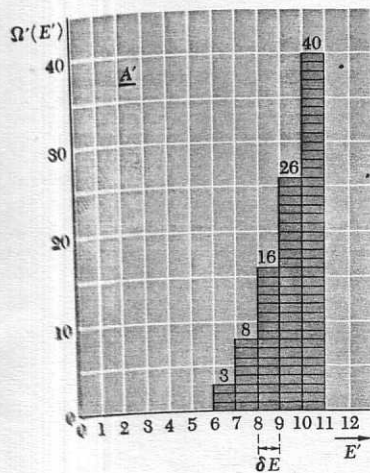
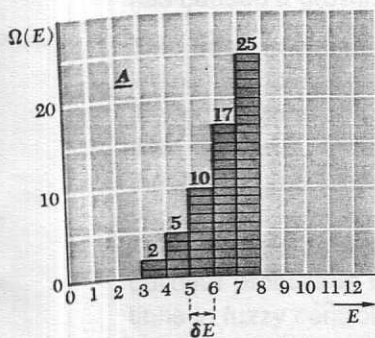


Fig. 4.1 Graphs showing, in the case of two special and very small systems A and A', the number of states $\Omega(E)$ accessible to A and the number of states $\Omega'(E')$ accessible to A' as a function of their respective energies E and E'. The energies are measured in terms of an arbitrary unit, only a few values of $\Omega(E)$ and $\Omega'(E')$ are shown.

equal to E? By the general argument leading to (3.20), the desired probability $P(E)$ then is simply given by

$$P(E) = \frac{\Omega^*(E)}{\Omega^*_{\text{tot}}} = C \Omega^*(E) \quad (3)$$

where $C = (\Omega^*_{\text{tot}})^{-1}$ is merely some constant independent of E.

The number $\Omega^*(E)$ can be readily expressed in terms of the numbers of states accessible to A and A', respectively. When A has an energy E, it can be in any one of its $\Omega(E)$ possible states. The system A' then must have an energy E' given by (2) by virtue of the conservation of energy. Hence A' can be in any one of the $\Omega'(E') = \Omega'(E^* - E)$ states accessible to it under these conditions. Since every possible state of A can be combined with every possible state of A' to give a different possible state of the total system A*, it follows that the number of distinct states accessible to A* when A has an energy E is simply given by the product

$$\Omega^*(E) = \Omega(E) \Omega'(E^* - E). \quad (4)$$

Correspondingly the probability (3) that the system A has an energy E is simply given by

$$P(E) = C \Omega(E) \Omega'(E^* - E). \quad (5)$$

Example

The following simple example uses very small numbers unrepresentative of real macroscopic systems, but serves to illustrate the essential ideas of the preceding paragraphs. Consider two special systems A and A' for which $\Omega(E)$ and $\Omega'(E')$ depend on their respective energies E and E' in the manner shown in Fig. 4.1. Here the energies E and E' are measured in terms of some arbitrary unit and are subdivided into unit energy intervals. Suppose that the combined energy E^* of both systems is equal to 13 units. A possible situation would be one where $E = 3$; it then follows that $E' = 10$. In

this case A can be in either one of its 2 possible states, and A' in any one of its 40 possible states. There are then a total of $\Omega^* = 2 \times 40 = 80$ different possible states accessible to the combined system A*. Table 4.1 enumerates systematically the possible situations compatible with the specified total energy E^* . Note that it would be most probable in a statistical ensemble of such systems to find the combined system A* in a state where $E = 5$ and $E' = 8$. This situation would be likely to occur twice as frequently as the situation where $E = 3$ and $E' = 10$.

(4) †

Let us now investigate the dependence of $P(E)$ on the energy E. Since A and A' are both systems with very many degrees of freedom, we know by (3.37) that $\Omega(E)$ and $\Omega'(E)$ are extremely rapidly increas-

ing functions of E and E' , respectively. Considering the expression (5) as a function of increasing energy E , it thus follows that the factor $\Omega(E)$ increases extremely rapidly while the factor $\Omega'(E^* - E)$ decreases extremely rapidly. The result is that the product of these two factors, i.e., the probability $P(E)$, exhibits a very sharp maximum† for some particular value $\bar{E}$ of the energy E . Thus the dependence of $P(E)$ on E must show the general behavior illustrated in Fig. 4.2 where the width ΔE of the region where $P(E)$ has appreciable magnitude is such that $\Delta E \ll \bar{E}$.

It is actually more convenient to investigate the behavior of $\ln P(E)$ rather than that of $P(E)$ itself since the logarithm is a much more slowly varying function of E . In addition, it follows by (5) that this logarithm involves the numbers Ω and Ω' as a simple sum rather than as a product, i.e.,

$$\ln P(E) = \ln C + \ln \Omega(E) + \ln \Omega'(E') \quad (6)$$

where $E' = E^* - E$. The value $E = \bar{E}$ which corresponds to the maximum of $\ln P(E)$ is determined by the condition‡

$$\frac{\partial \ln P}{\partial E} = \frac{1}{P} \frac{\partial P}{\partial E} = 0 \quad (7)$$

and thus corresponds also to the maximum of $P(E)$ itself. By using (6) and (2), the condition (7) becomes simply

$$\frac{\partial \ln \Omega(E)}{\partial E} + \frac{\partial \ln \Omega'(E')}{\partial E'} (-1) = 0$$

or

$$\beta(E) = \beta'(E') \quad (8)$$

where we have introduced the definition

$$\beta(E) \equiv \frac{\partial \ln \Omega}{\partial E} = \frac{1}{\Omega} \frac{\partial \Omega}{\partial E} \quad (9)$$

and the corresponding definition for $\beta'(E')$. The relation (8) is thus the fundamental condition which determines the particular value $\bar{E}$ of the energy of A (and corresponding value $\bar{E}' \equiv E^* - \bar{E}$ of the energy of A') which occurs with the greatest probability $P(E)$.

E	E'	$\Omega(E)$	$\Omega'(E')$	$\Omega^*(E)$
3	10	2	40	80
4	9	5	26	130
5	8	10	16	160
6	7	17	8	136
7	6	25	3	75

Table 4.1 Enumeration of the possible numbers of states compatible with a specified total energy $E^* = 13$ of the systems A and A' described in Fig. 4.1.

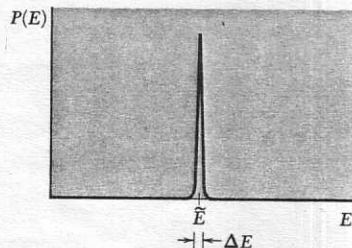


Fig. 4.2 Schematic illustration showing the dependence of the probability $P(E)$ on the energy E .

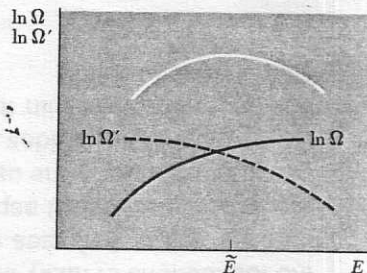


Fig. 4.3 Schematic sketch of the dependence of $\ln \Omega(E)$ and of $\ln \Omega'(E') \equiv \ln \Omega'(E^* - E)$ on the energy E . By virtue of (3.38), the energy dependence of $\ln \Omega$ is roughly of the form $\ln \Omega(E) \sim f \ln (E - E_0) + \text{constant}$. Since the curves are thus concave downward, their sum (shown in white) exhibits a unique maximum at some value $\bar{E}$. This gentle maximum of the slowly varying logarithm of $P(E)$, given by (6), corresponds then to an extremely sharp maximum of $P(E)$ itself.

† Note that the behavior of $P(E)$ is analogous to that of the preceding simple example, except for the fact that the maximum of $P(E)$ is enormously sharper for macroscopic systems where $\Omega(E)$ and $\Omega'(E')$ are such rapidly varying functions.

‡ We write this as a partial derivative to emphasize that all external parameters of the system are supposed to remain fixed throughout the discussion.

Sharpness of the maximum of $P(E)$

By examining the behavior of $\ln P(E)$ near its maximum, one can readily estimate how rapidly $P(E)$ decreases when E becomes different from $\bar{E}$. Indeed, it is shown in Appendix A.3 that $P(E)$ becomes negligibly small compared to its maximum value when E differs from $\bar{E}$ by appreciably more than an amount ΔE whose magnitude is roughly of the order

$$\Delta E \sim \frac{\bar{E}}{\sqrt{f}}. \quad (10)$$

Here f is the number of degrees of freedom of the smaller of the two interacting systems and $\bar{E}$ is assumed to be much greater than the lowest (or ground state) energy of A . For a typical system consisting of a mole of atoms, f is of the

order of Avogadro's number so that $f \sim 10^{24}$ and

$$\Delta E \sim 10^{-12} \bar{E}. \quad (11)$$

Hence the probability $P(E)$ has ordinarily an exceedingly sharp maximum at the value $\bar{E}$, becoming negligibly small when E differs from $\bar{E}$ by as little as a few parts in 10^{12} . Thus the energy of A practically never differs appreciably from $\bar{E}$; in particular, the mean value $\bar{E}$ of A must then also be equal to $\bar{E}$, i.e., $E = \bar{E}$. Here we encounter another instance where the relative magnitude of the fluctuations of a quantity about its mean value is exceedingly small when one is dealing with a system consisting of very many particles.

Some conventional definitions

The preceding discussion shows that the quantities $\ln \Omega$ and β of the system A (and the corresponding quantities of A') are of crucial importance in the study of thermal interaction. It is, therefore, convenient to introduce some alternative symbols and conventional names for these quantities.

Note first that the parameter β has, according to its definition (9), the dimensions of a reciprocal energy. It is thus often useful to express β^{-1} as a multiple of some positive constant k which has the dimensions of energy and thus can be expressed in ergs. (This constant is called *Boltzmann's constant* and its magnitude can be chosen once and for all in some convenient way.) The parameter β^{-1} can then be written in the form

$$\boxed{\frac{1}{\beta} \equiv kT} \quad (12)$$

+ where the quantity T thus defined provides a measure of energy in units of the quantity k . This new parameter T is called the *absolute temperature* of the system under consideration and its magnitude is commonly expressed in terms of *degrees*.† The physical reason for the name "temperature" will become more apparent in Sec. 4.3.

† For example, an absolute temperature of 5 degrees corresponds merely to an energy of $5k$.

By virtue of (9), the definition of T in terms of $\ln \Omega$ can now be written in the form

$$\frac{1}{T} \equiv \frac{\partial S}{\partial E} \quad (13)$$

where we have introduced the quantity S defined by

$$S \equiv k \ln \Omega. \quad (14)$$

This quantity S is called the *entropy* of the system under consideration.

- It has the dimensions of energy because its definition involves the constant k . According to its definition (14), the entropy of a system is merely a logarithmic measure of the number of states accessible to the system. In accordance with the comments at the end of Sec. 3.6, the
- † entropy thus provides a quantitative measure of the degree of randomness of the system.†

- In terms of the previous definitions, the condition that the probability $P(E)$ be maximum is, by virtue of (3), equivalent to the statement that the entropy $S^* \equiv k \ln \Omega^*$ of the total system is maximum with respect to the energy E of the subsystem A . Using (6), the condition of maximum probability is thus equivalent to the statement that

$$S^* = S + S' = \text{maximum}. \quad (15)$$

- This condition is fulfilled if (8) is satisfied, i.e., if

$$T = T'. \quad (16)$$

- † Our discussion thus shows that the energy E of A adjusts itself in such a way that the entropy of the total isolated system A^* is as large as possible. The system A^* is then distributed over the largest possible number of states, i.e., it is in its most random macrostate.

4.2 The Approach to Thermal Equilibrium

- As we have seen, the probability $P(E)$ has an exceedingly sharp maximum at the energy $E = \bar{E}$. Therefore, in the equilibrium situation where A and A' are in thermal contact, the system A almost always has an energy E extremely close to $\bar{E}$, while the system A' correspondingly has an energy E' extremely close to $\bar{E}' = E^* - \bar{E}$. To excellent ap-

† Note that the entropy defined by (14) has a definite value which is, by virtue of (3.40), essentially independent of the magnitude of the energy subdivision interval δE used in our discussion. Furthermore, since δE is some fixed interval independent of E , the derivative (9) defining β or T is certainly also independent of δE .

proximation, the respective mean energies of the systems are then also equal to these energies, i.e.,

$$E = \bar{E} \quad \text{and} \quad E' = \bar{E}'. \quad (17)$$

Now consider a situation where the systems A and A' are initially isolated from each other and separately in equilibrium, their mean energies being E_i and E'_i , respectively. Suppose that A and A' are then placed in thermal contact so that they are free to exchange energy with each other. The resulting situation is ordinarily an extremely improbable one, except in the special case where the respective energies of the systems were initially very close to $\bar{E}$ and $\bar{E}'$. In accordance with our postulate (3.18), the systems will then tend to exchange energy until they attain the ultimate equilibrium situation discussed in the preceding section. The final mean energies E_f and E'_f of the systems will then, by (17), be equal to

$$E_f = \bar{E} \quad \text{and} \quad E'_f = \bar{E}' \quad (18)$$

so that the probability $P(E)$ becomes maximum. The β parameters of the systems will then correspondingly be equal, i.e.,

$$\beta_f = \beta'_f \quad (19)$$

where $\beta_f \equiv \beta(E_f)$ and $\beta'_f \equiv \beta(E'_f)$.

The conclusion that the systems exchange energy until they reach the situation of maximum probability $P(E)$ is, by virtue of (6) and the definition (14), equivalent to the statement that they exchange energy until their total entropy becomes maximum. Thus the final probability (or entropy) can never be less than the original one, i.e.,

$$S(E_f) + S'(E'_f) \geq S(E_i) + S'(E'_i)$$

or

$$\Delta S + \Delta S' \geq 0, \quad (20)$$

where

$$\Delta S \equiv S(E_f) - S(E_i)$$

and

$$\Delta S' \equiv S'(E'_f) - S'(E'_i)$$

denote the entropy changes of A and A' , respectively.

In the process of exchanging energy, the total energy of the systems is, of course, conserved. In accordance with (3.49) and (3.50), it thus follows that

$$Q + Q' = 0 \quad (21)$$

where Q and Q' denote the heats absorbed by A and A' respectively.

† The relations (20) and (21) summarize completely the conditions which must be satisfied in any process of thermal interaction.

.. Our discussion shows that there are two possible cases which may arise:

(i) The initial energies of the systems are such that $\beta_i = \beta_i'$, where $\beta_i \equiv \beta(E_i)$ and $\beta_i' \equiv \beta(E_i')$. The systems are then already in their most probable situation, i.e., their total entropy is already maximum. Hence the systems remain in equilibrium and no net exchange of heat occurs between them.

(ii) More generally, the initial energies of the systems are such that $\beta_i \neq \beta_i'$. The systems are then in a rather improbable situation where their total entropy is not maximum. Hence the situation will change in time while energy, in the form of heat, is exchanged between the systems until the final equilibrium is reached where the total entropy is maximum and $\beta_f = \beta_f'$.

4.3 Temperature

In the last section, we noted that the parameter β [or equivalently the parameter $T = (k\beta)^{-1}$] has the following two properties:

(i) If two systems separately in equilibrium are characterized by the *same* value of the parameter, then equilibrium will be preserved and *no* heat transfer will occur when the systems are brought into thermal contact with each other.

(ii) If the systems are characterized by *different* values of the parameter, then equilibrium will not be preserved and heat transfer *will* occur when the systems are brought into thermal contact with each other.

.. These statements allow us to derive several important consequences. In particular, they permit us to formulate in a precise and quantitative way the qualitative notions explored in Sec. 1.5.

Imagine, for example, that there are three systems A, B, and C which are separately in equilibrium. Suppose that no heat transfer occurs when C is placed in thermal contact with A, and that no heat transfer occurs when C is placed in thermal contact with B. Then we know that $\beta_C = \beta_A$ and that $\beta_C = \beta_B$ (where β_A , β_B , and β_C denote the β parameters of systems A, B, and C, respectively). But from these two equalities we can conclude that $\beta_A = \beta_B$, and hence that no heat transfer will occur if systems A and B are placed in thermal contact with each other.

Thus we are led to the following general conclusion:

If two systems are in thermal equilibrium with a third system, then they must be in thermal equilibrium with each other. (22)

The statement (22) is called the *zeroth law of thermodynamics*. Its validity makes possible the use of test systems, called *thermometers*, which allow measurements to decide whether any two systems will, or will not, exchange heat when brought into thermal contact with each other. Such a thermometer can be *any* macroscopic system M chosen in accordance with the following two specifications:

(a) Among the many macroscopic parameters characterizing the system M , select one (call it θ) which varies by appreciable amounts when M gains or loses energy by thermal interaction. All the other macroscopic parameters of M are held fixed. The parameter θ which is allowed to vary is called the *thermometric parameter* of M .

(b) The system M is ordinarily chosen to be much smaller (i.e., to have many fewer degrees of freedom) than the systems which it is designed to test. This is desirable in order to minimize any energy transfer to these systems and thus to reduce to a minimum their disturbance by the testing process.

Examples of thermometers

There are many possible systems which can serve as thermometers. We mention only a few of those which are commonly used.

(i) A liquid, such as mercury or alcohol, confined within a glass tube of small diameter. This is the familiar type of thermometer which we have already described in Sec. 1.5. Here the thermometric parameter θ is the height of the liquid in the tube.

(ii) A gas confined within a bulb in such a way that its volume is kept constant. This is called a *constant-volume gas thermometer*. Here the thermometric parameter θ is the pressure exerted by the gas.

(iii) A gas confined within a bulb in such a way that its pressure is kept constant. This is called a *constant-pressure*

gas thermometer. Here the thermometric parameter θ is the volume occupied by the gas.

(iv) An electrical conductor (e.g., a coil of platinum wire) maintained at constant pressure and carrying a small current. This is called a *resistance thermometer*. Here the thermometric parameter θ is the electrical resistance of the conductor.

(v) A paramagnetic sample maintained at constant pressure. Here the thermometric parameter θ is the magnetic susceptibility of the sample (i.e., the ratio of its mean magnetic moment per unit volume to the applied magnetic field.) This quantity can be determined, for example, by measuring the self-inductance of a coil wound around the sample.

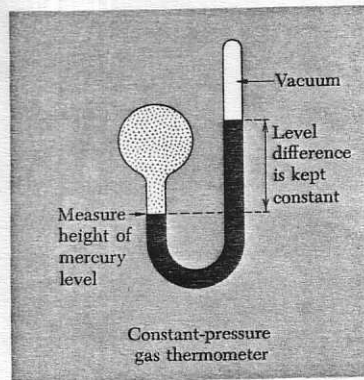
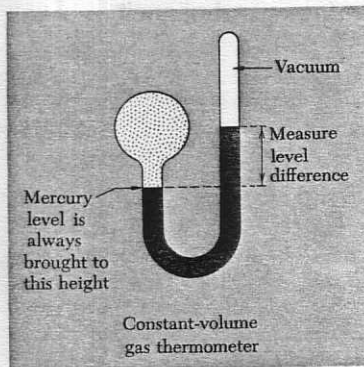


Fig. 4.4 Constant-volume and constant-pressure gas thermometers.

$O \in E \rightarrow$ For O to observe $E \Rightarrow O$ actually observes what is in the neighbourhood to $\{\delta E\} \ni O \approx M$, generally speaking $O \in \{E\}$.
Of course, for O to observe $\{E\}$ either always changes $\{\delta E\}$ and this is done by different O 's.

* Another method for observing $\{E\}$ is for O to become $\{E\}$
 $\Rightarrow M \approx O \approx \{E\}$, if this is not possible for O , $\rightarrow O$ must invent G which collects for O all the observations $\{E\}$

* also $O_1 \in \delta E_1$ can go to $O_2 \in \delta E_2$
to ask about δE_2 for $O \approx \{E\} \rightarrow O(G)$, $G \equiv \{E\}$

* These ideas can find their use in nonequilibrium systems,
general system behavior, generation of E , $G \ni \{E\}$

" O is contained in E and E belongs to G ."

- (2) A thermometer M is used in the following way. It is successively placed in thermal contact with the systems under test, call them A and B , and is allowed to come to equilibrium with each.

(i) Suppose that the thermometric parameter θ of M (e.g., the length of the liquid column of a mercury-in-glass thermometer) assumes the same value in both cases. This implies that, after M has come to equilibrium with A , it remains in equilibrium after being placed in thermal contact with B . Hence the zeroth law allows one to conclude that A and B will remain in equilibrium if they are brought into thermal contact with each other.

(ii) Suppose that the thermometric parameter θ of M does *not* assume the same value in both cases. Then one can conclude that A and B will *not* remain in equilibrium if they are brought into thermal contact with each other. To make the argument explicit, suppose that A and B *did* remain in equilibrium. After M attains thermal equilibrium with A , it would then, according to the zeroth law, have to remain in equilibrium when brought into thermal contact with B . But then the parameter θ cannot change if M is brought into thermal contact with B , contrary to hypothesis.†

Consider *any* thermometer M with *any one* parameter θ chosen as its thermometric parameter. The value assumed by θ when the thermometer M has come to thermal equilibrium with some system A will, by definition, be called the *temperature of A with respect to the particular thermometric parameter θ of the particular thermometer M* . According to this definition the temperature can be a length, a pressure, or any other quantity. Note that, even if two different thermometers have parameters of the same kind, they do not ordinarily yield the same value of the temperature for the same body.‡ Furthermore, if a body C has a temperature halfway between the temperatures of bodies A and B when measured by one thermometer, this statement is not necessarily true with respect to the temperatures measured by some other thermometer. Nevertheless, our discussion shows that the tem-

† All the preceding measurements could be performed with any other thermometer M' having a thermometric parameter θ' . There exists ordinarily a one-to-one relation between any value of θ and the corresponding value of θ' . In exceptional cases a particular thermometer M might, however, be multiple-valued so that a given value of θ corresponds to more than one value of θ' for almost every other thermometer M' . Peculiar thermometers which are multiple-valued in the experimental range of interest are rarely useful and will be excluded from our discussion. (See Prob. 4.1.)

‡ For example, the two thermometers might both consist of glass tubes filled with liquid, so that the length of the liquid column is the thermometric parameter in both cases. The liquid might, however, be mercury in one thermometer and alcohol in the other.

perature concept which we have defined has the following useful property:

Two systems will remain in equilibrium when placed in thermal contact with each other if, and only if, they have the same temperature with respect to the same thermometer. (23)

† The temperature concept which we have introduced is important and useful, but is rather arbitrary in the sense that the temperature assigned to a system depends in an essential way on the peculiar properties of the particular system M used as the thermometer. On the other hand, we can exploit the properties of the parameter β to obtain

(i) a much more useful temperature. Indeed, suppose that we have a thermometer M whose parameter β is known as a function of its thermometric parameter θ . If this thermometer is placed in thermal contact with some system A , then we know that in equilibrium $\beta = \beta_A$.

Thus the thermometer measures, by virtue of (9), a fundamental property of the system A , namely the fractional increase of the number of its states with energy. Furthermore, suppose that we take any other thermometer M' whose parameter β' is also known as a function of its thermometric parameter θ' . If this thermometer is placed in thermal contact with the system A , then we know that in equilibrium $\beta' = \beta_A$. Thus $\beta' = \beta$ and we arrive at the following conclusion:

(ii) (a) If the parameter β is used as the thermometric parameter of a thermometer, then every such thermometer yields the same temperature reading when used to measure the temperature of a particular system. Furthermore, this temperature measures a fundamental property of the number of states of the system under test. (24)

‡ The parameter β is, therefore, a particularly useful and fundamental temperature parameter. This is the reason for the name of *absolute temperature* given to the corresponding temperature parameter $T \equiv (k\beta)^{-1}$ defined in terms of β . We shall postpone until the next chapter a discussion of the following two points: (i) practical procedures for finding numerical values of β or T by appropriate measurements, and (ii) the international convention adopted to assign a particular numerical value to k .

$$\Omega \equiv \Omega(E)$$

$$\Omega(d\Omega/dE) \equiv \frac{1}{k} = kT$$

Properties of the absolute temperature

By virtue of its definition (9), the absolute temperature is given by

$$\frac{1}{kT} \equiv \beta \equiv \frac{\partial \ln \Omega}{\partial E}. \quad (25)$$

We saw in (3.37) that $\Omega(E)$ is, for any ordinary system, an extremely rapidly increasing function of its energy E . Hence (25) implies that,

$$\text{for any ordinary system,} \quad \beta > 0 \quad \text{or} \quad T > 0. \quad (26)$$

In other words,

the absolute temperature of any ordinary system is positive.†

(27)

We can also readily estimate the order of magnitude of the absolute temperature of a system. The approximate functional dependence of $\Omega(E)$ is ordinarily of the form given by (3.38),

$$\Omega(E) \propto (E - E_0)^f \quad (28)$$

where f is the number of degrees of freedom of the system and E is the energy of the system whose ground state energy is E_0 . Thus

$$\ln \Omega \sim f \ln (E - E_0) + \text{constant}$$

so that

$$\beta = \frac{\partial \ln \Omega}{\partial E} \sim \frac{f}{E - E_0}. \quad (29)$$

The magnitude of T can then be estimated by putting $E = \bar{E}$, the mean energy of the system. Hence we can conclude that,

$$\text{for an ordinary system,} \quad kT = \frac{1}{\beta} \sim \frac{\bar{E} - E_0}{f}. \quad (30)$$

In other words,

for any ordinary system at the absolute temperature T , the quantity kT is roughly equal to the mean energy (above the ground state) per degree of freedom of the system.

(31)

† As pointed out in connection with (3.38), the qualifying phrase *for any ordinary system* is meant to exclude specifically the exceptional case of a system where the kinetic energy of the particles is ignored and where their spins have a magnetic energy which is sufficiently high.

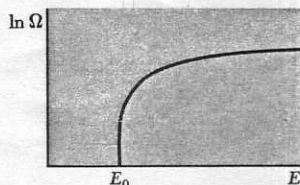


Fig. 4.5 Schematic sketch showing the behavior of $\ln \Omega$ as a function of the energy E . The slope of the curve is the absolute temperature parameter β .

Given a system A for which $E = E_0 \rightarrow E_0$ $\rightarrow T = 0$ which implies that each atom (molecule) stops from doing any vibration i.e. stops motion and become fixed in the position of equilibrium. $* f = \sum_{i=1}^N f_i$; f_i is the number of degrees of freedom of the fixed atom or molecule.

The condition of equilibrium (8) between two systems in thermal contact asserts that their respective absolute temperatures must be equal. By virtue of (31) we see that this condition is roughly equivalent to the statement that the total energy of the interacting systems is shared between them in such a way that the mean energy per degree of freedom is the same for both systems. This last statement was essentially the one which we used in our qualitative discussion of Sec. 1.5.

How does the parameter β , or T , vary with the energy E of a system? The quantity β measures the slope of the curve of $\ln \Omega$ versus E . We already noted in the caption of Fig. 4.3 that this curve must be concave downward so as to guarantee the physical requirement that a situation with a unique maximum probability results when two systems are placed in thermal contact with each other. Hence it follows that the slope of the curve must decrease as E increases; i.e.,

$$\text{for any system,} \quad \frac{\partial \beta}{\partial E} < 0. \quad (32)$$

In the case of an ordinary system, this result follows also from the approximate functional form (28) since differentiation of (29) yields explicitly

$$\frac{\partial \beta}{\partial E} \sim -\frac{f}{(E - E_0)^2} < 0. \quad (33)$$

Since we have just shown that β decreases as E increases and since, by its definition $T \equiv (k\beta)^{-1}$, the absolute temperature T increases as β decreases, we can then conclude from (32) that

$$\boxed{\text{the absolute temperature of any system is an increasing function of its energy.}} \quad (34)$$

In more mathematical terms,

$$\frac{\partial T}{\partial E} = \frac{\partial}{\partial E} \left(\frac{1}{k\beta} \right) = -\frac{1}{k\beta^2} \frac{\partial \beta}{\partial E}$$

so that (32) implies

$$\frac{\partial T}{\partial E} > 0. \quad (35)$$

† This last relation permits us to establish a general connection between absolute temperatures and the direction of heat flow. Consider two systems A and A' which are initially in equilibrium at different absolute temperatures T_i and T_i' , and which are then brought into thermal contact with each other. One system then absorbs heat, and the other gives off heat, until the systems attain a final equilibrium at some common absolute temperature T_f . Suppose that system A is the one which absorbs heat and thus gains energy; then it follows by (34) that $T_f > T_i$. Correspondingly system A' must give off heat and thus lose energy; then it follows by (34) that $T_f < T_i'$. Hence the initial and final temperatures are such that

$$T_i < T_f < T_i'.$$

“ This means that the system A which absorbs heat has an initial absolute temperature T_i which is less than the initial absolute temperature T_i' of the system A' which gives off heat. In short,

when any two ordinary systems are placed in thermal contact with each other, heat is given off by the system with the larger absolute temperature and is absorbed by the system with the smaller absolute temperature.†

(36)

• Temperature ~ Information
• Capability of absorption
• depends on the system
• which of who absorbs.

Since we defined the *warmer* system as the one which gives off heat and the *colder* system as the one which absorbs heat, (36) is equivalent to the statement that *a warmer system has a higher absolute temperature than a colder one*.

4.4 Small Heat Transfer

• The preceding sections complete our general discussion of the thermal interaction between macroscopic systems. We shall now turn our attention to some special simple cases which are of particular importance.

Suppose that, when a system A is placed in thermal contact with some other system, it absorbs an amount of heat Q which is so small that

$$|Q| \ll E - E_0, \quad (37)$$

i.e., so that the resulting mean energy change $\Delta E = Q$ of A is small compared to the mean energy E of A above its ground state. The

† In the exceptional case of a system of spins, this statement may need to be qualified because $T \rightarrow \pm\infty$ when $\beta \rightarrow 0$. See Prob. 4.30.

absolute temperature of the system A changes then by a negligible amount. Indeed, putting $E = \bar{E}$, (29) and (33) yield the estimates

$$\lim_{Q \rightarrow 0} \frac{\Delta \beta}{\Delta \bar{E}} \Delta \bar{E} = \Delta \beta = \frac{\partial \beta}{\partial E} Q \sim -\frac{f}{(E - E_0)^2} Q \sim -\frac{\beta}{E - E_0} Q.$$

Hence (37) implies that

$$|\Delta \beta| = \left| \frac{\partial \beta}{\partial E} Q \right| \ll \beta. \quad (38)$$

Since $T = (k\beta)^{-1}$ or $\ln T = -\ln \beta - \ln k$, it follows correspondingly that $(\Delta T/T) = -(\Delta \beta/\beta)$ so that (38) is also equivalent to

$$|\Delta T| \ll T. \quad (39)$$

We shall say that the heat Q absorbed by a system is *small* whenever (38) is valid, i.e., whenever Q is sufficiently small so that the absolute temperature of the system remains essentially unchanged.

Suppose that the system A absorbs such a small amount of heat Q . Its initial and final energies are then, with overwhelming probability, equal to their respective mean values $\bar{E}$ and $\bar{E} + Q$. In the process of absorbing this heat, the number of states $\Omega(E)$ accessible to A also changes. By expanding in a Taylor series, we find

$$\begin{aligned} \ln \Omega(\bar{E} + Q) - \ln \Omega(\bar{E}) &= \left(\frac{\partial \ln \Omega}{\partial E} \right) Q + \frac{1}{2} \left(\frac{\partial^2 \ln \Omega}{\partial E^2} \right) Q^2 + \dots \\ &= \beta Q + \frac{1}{2} \frac{\partial \beta}{\partial E} Q^2 + \dots \end{aligned}$$

But since the absorbed heat Q is supposed to be small, the absolute temperature of A remains essentially unchanged. Hence the term involving $\partial \beta / \partial E$ can be neglected, in accordance with (38). The change in the quantity $\ln \Omega$ then becomes simply

$$\Delta(\ln \Omega) = \frac{\partial \ln \Omega}{\partial E} Q = \beta Q. \quad (40)$$

In the process of absorbing an amount of heat Q , the entropy $S \equiv k \ln \Omega$ of a system at the absolute temperature $T = (k\beta)^{-1}$ changes thus by an amount ΔS which is,

if Q is small,

$$\Delta S = \frac{Q}{T}.$$

(41)

We emphasize that, even if the heat Q is large in absolute magnitude, it may yet be relatively small in the comparative sense of (37) or (39) so that the relation (41) remains valid. If the absorbed heat is actually infinitesimal in magnitude, we can denote it by dQ ; the corresponding infinitesimal entropy change is then equal to

$$dS = \frac{dQ}{T}. \quad (42)$$

Note that the heat dQ is merely an infinitesimal quantity. The quantity dS is, however, an actual differential, i.e., an infinitesimal *difference* between the entropies of A in its final and initial macrostates.

The heat Q absorbed by a system A will always be very small in the comparative sense of (37) or (39) when A is placed in thermal contact with any other system B sufficiently smaller than A itself. Indeed, any amount of heat Q which A can absorb from B then is at most of the order of the total energy of B (above its ground state) and is thus much smaller than the energy $E - E_0$ of A itself. A system A is said to act as a *heat reservoir* (or *heat bath*) with respect to some other set of systems if it is sufficiently large so that its temperature remains essentially unchanged in any thermal interaction with those systems. Equation (41) thus is always valid in relating the entropy change ΔS of a heat reservoir to the heat Q absorbed by it.

dQ is infinitesimal because it is a transferred (absorbed) quantity and hence cannot be a function of some variable, it is S , of the system affected by dQ , that is a function and therefore can be differentiated.

4.5 System in Contact with a Heat Reservoir

Most systems which one encounters in practice are not isolated, but are free to exchange heat with their environment. Since such a system is usually small compared to its environment, it constitutes a relatively small system in thermal contact with a heat reservoir provided by other systems in its environment. (For example, any object in a room, say a table, is in thermal contact with the heat reservoir consisting of the room itself with its floor, walls, other furniture, and enclosed air.) In this section we shall, therefore, consider any relatively small system A in contact with a heat reservoir A' , and ask the following detailed question about the small system A : Under conditions of equilibrium, what is the probability P_r of finding the system A in any one particular state r of energy E_r ?

This is a rather important and very general kind of question. Note that in the present context the system A can be any system which has many fewer degrees of freedom than the heat reservoir A' . Thus A

might be any relatively small *macroscopic* system. (For example, it might be a piece of copper immersed in the water of a lake, the latter acting as a heat reservoir.) Alternatively A might also be a distinguishable *microscopic* system which can be clearly identified.† (For example, it might be an atom located at a particular lattice site in a solid, the latter acting as a heat reservoir.)

To facilitate counting the states of the reservoir A' , we again imagine its energy scale to be subdivided into small fixed intervals of magnitude δE and denote by $\Omega'(E')$ the number of states accessible to A' when its energy is equal to E' (i.e., when it lies between E' and $E' + \delta E$). (Here δE is assumed to be very small compared to the separation between the energy levels of A , but to be large enough to contain many possible states of the reservoir A' .) It is then quite easy to use reasoning similar to that of Sec. 4.1 to find the desired probability P_r that the system A is in its state r . Although the reservoir can have any energy E' , the conservation of energy applied to the isolated system A^* consisting of both A and A' requires that the energy of A^* have some constant value, say E^* . When the system A is in its state r of energy E_r , the reservoir A' must then have an energy

$$E' = E^* - E_r. \quad (43)$$

But, when A is in this one definite state r , the number of states accessible to the combined system A^* is simply the number of states $\Omega'(E^* - E_r)$ which are accessible to A' . Our fundamental statistical postulate asserts, however, that the isolated system A^* is equally likely to be found in each one of its accessible states. Hence the probability of occurrence of a situation where A is in the state r is simply proportional to the number of states accessible to A^* when A is known to be in the state r , i.e.,

$$P_r \propto \Omega'(E^* - E_r). \quad (44)$$

Up to this point our discussion has been completely general. Let us now make use of the fact that A is very much smaller than the reservoir A' in the sense that any energy E_r of interest to us satisfies the relation

$$E_r \ll E^*. \quad (45)$$

† This qualifying remark is pertinent since it may not always be possible, in a quantum-mechanical description, to identify a particular particle among particles which are basically indistinguishable.

$$P_r(E_r) \propto \Omega^*(E_r)$$

$$\Omega^*(E_r) = 1 \times \Omega'(E^* - E_r)$$

We can then find an excellent approximation for (44) by expanding the slowly varying logarithm of $\Omega'(E')$ about the value $E' = E^*$. Analogously to (40), we thus get for the heat reservoir A'

$$\begin{aligned}\ln \Omega'(E^* - E_r) &= \ln \Omega'(E^*) - \left[\frac{\partial \ln \Omega'}{\partial E'} \right] E_r \\ &= \ln \Omega'(E^*) - \beta E_r.\end{aligned}\quad (46)$$

Here we have written

$$\beta \equiv \left[\frac{\partial \ln \Omega'}{\partial E'} \right] \approx \frac{1}{kT} \quad (47)$$

for the derivative evaluated at the fixed energy $E' = E^*$. Thus $\beta = (kT)^{-1}$ is simply the constant temperature parameter of the *heat reservoir* A' .† Hence (46) yields the result

$$P_r \propto \Omega'(E^* - E_r) = \Omega'(E^*) e^{-\beta E_r}. \quad (48)$$

Since $\Omega'(E^*)$ is merely a constant independent of r , (44) then becomes simply

$$P_r = C e^{-\beta E_r} \quad (49)$$

where C is a constant of proportionality independent of r .

Let us examine the physical content of the results (44) or (49). If A is known to be in a definite state r , the reservoir A' can be in any one of the large number $\Omega'(E^* - E_r)$ of states accessible to it under these conditions. But the number of states $\Omega'(E')$ accessible to the reservoir is ordinarily a rapidly increasing function of its energy E' [i.e., β in (47) is ordinarily positive]. Suppose then that we compare the probabilities of finding the system A in *any* two of its states which have different energies. If A is in the state where its energy is higher, the conservation of energy for the total system then implies that the energy of the reservoir is correspondingly lower; hence the number of states accessible to the reservoir is markedly reduced. Accordingly, the probability of encountering this situation is very much less. The exponential dependence of P_r on E_r in (49) merely expresses the result of this argument in mathematical terms.

† For the sake of simplicity, we shall not embellish the symbol β with a prime.

$e^{-\beta E_r}$ decreases as (T) nonincreasing
this happens at a far
distance from (A) .

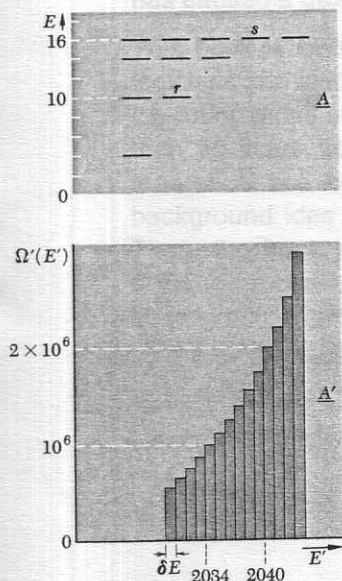


Fig. 4.6 Schematic illustration showing the states accessible to a particular system A and to a special (rather small) heat reservoir A'. The top diagram shows the energy levels corresponding to a few distinct states of A. The bottom diagram shows, for a few values of E' , the number of states $\Omega'(E')$ accessible to A' as a function of its energy E' . The energy is measured in terms of an arbitrary unit.

Example

Let us illustrate the preceding comments with a simple example. Consider a system A which has energy levels some of which are shown in the top part of Fig. 4.6. Consider also the much larger system A' whose energy scale is subdivided into intervals of size $\delta E = 1$ unit and whose number of states $\Omega'(E')$ is shown as a function of its energy E' in the bottom part of Fig. 4.6. We assume that A is in thermal equilibrium with the heat reservoir A' and that the total energy E^* of the combined system A* has some value $E^* = 2050$ units. Suppose that A is in the particular state r where its energy $E_r = 10$ units. The energy of the reservoir A' must then be $E' = 2040$; accordingly, A' can be in any one of

2×10^6 possible states. In an ensemble containing many isolated systems A* consisting of A and A', the number of instances where A is found in the state r will then be proportional to 2×10^6 . On the other hand, suppose that A is in the particular state s where its energy $E_s = 16$ units. The energy of the reservoir A' must then be $E' = 2034$; accordingly, A' can now be in any one of only 10^6 possible states. In the ensemble of systems, the number of instances where A is found in the state s will thus be proportional to 10^6 and will consequently be only half as large as the number of instances where A is in the state r of lower energy.

The probability (49) is a very general result of fundamental importance in statistical mechanics. The exponential factor $e^{-\beta E_r}$ is called the *Boltzmann factor*; the corresponding probability distribution (49) is known as the *canonical distribution*. An ensemble of systems all of which are in contact with a heat reservoir of known temperature T [i.e., all of which are distributed over their states in accordance with (49)] is called a *canonical ensemble*.

The constant of proportionality C in (49) can be readily determined by the normalization condition that the system must have probability unity of being in some one of its states; i.e.,

$$\sum_r P_r = 1 \quad (50)$$

where the sum extends over all possible states of A irrespective of energy. By virtue of (49) this condition determines C to be such that

$$C \sum_r e^{-\beta E_r} = 1.$$

Hence (49) can be written in the explicit form

$$P_r = \frac{e^{-\beta E_r}}{\sum_r e^{-\beta E_r}}. \quad (51)$$

- The probability distribution (49) allows us to calculate very simply the mean values of various parameters characterizing the system A in contact with a heat reservoir at the absolute temperature $T = (k\beta)^{-1}$.
- For example, let y be any quantity assuming the value y_r in state r of the system A . Then the mean value of y is given by

$$\bar{y} \equiv \sum_r P_r y_r = \frac{\sum_r e^{-\beta E_r} y_r}{\sum_r e^{-\beta E_r}} \quad (52)$$

where the summation is over all states r of the system A .

† Remarks pertinent when A is a macroscopic system

The fundamental result (49) gives the probability P_r of finding A in any one particular state r of energy E_r . The probability $P(E)$ that A has an energy in a small range, say between E and $E + \delta E$, is then simply obtained by adding the probabilities for all states r whose energy E_r lies in the range $E < E_r < E + \delta E$; i.e.,

$$P(E) = \sum_r P_r$$

where the prime on the summation symbol indicates that the sum extends only over the states having nearly identical energies lying in this small range. But the probability P_r is then, by (49), essentially the same for all these states and proportional to $e^{-\beta E}$. Hence the probability $P(E)$ of interest is obtained simply by multiplying (the probability of finding A in any one of these states) by (the number $\Omega(E)$ of its states in this energy range); i.e.,

$$P(E) = C \Omega(E) e^{-\beta E}. \quad (53)$$

To the extent that A itself is a large system (although very much smaller than A'), $\Omega(E)$ is a rapidly increasing function of E . The presence of the rapidly decreasing factor $e^{-\beta E}$ in (53) results then in a maximum of the product $\Omega(E) e^{-\beta E}$. This maximum in $P(E)$ is sharper the larger A is, i.e., the more rapidly $\Omega(E)$ increases with E . Thus we arrive again at the conclusions obtained in Sec. 4.1 for a macroscopic system.

When a system in contact with a heat reservoir is itself of macroscopic size, the relative magnitude of the fluctuations of its energy E is so exceedingly small that its energy is practically always equal to its mean value $\bar{E}$. On the other hand, if the system were removed from contact with the heat reservoir and were thermally isolated, its energy could not fluctuate at all. The distinction between this situation and the previous one, however, is so small that it is practically irrelevant; in particular, the mean values of all physical parameters of the system (such as its mean pressure or mean magnetic moment) remain quite unaffected. Hence it makes no difference whether these mean values are calculated by considering the macroscopic system to be an isolated system with a fixed energy in some small range between E and $E + \delta E$, or by considering this system to be in thermal contact with a heat reservoir of such a temperature that the mean energy $\bar{E}$ of the system is equal to E . The latter point of view, however, makes calculations much easier. The reason is that use of the canonical distribution reduces the computation of mean value to the evaluation of the sums (52) over all states without restriction; thus it does *not* require the much more difficult task of counting the number Ω of states of a particular type lying in a specified small energy range.

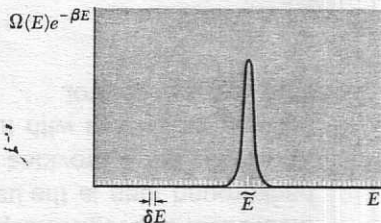
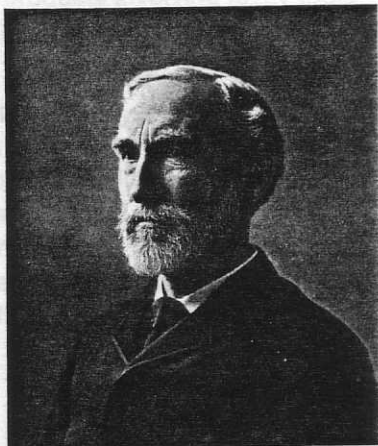


Fig. 4.7 Schematic illustration showing the dependence of the function $\Omega(E)e^{-\beta E}$ on the energy E of a macroscopic system in contact with a heat reservoir.



Ludwig Boltzmann

Fig. 4.8 Ludwig Boltzmann (1844–1906). This Austrian physicist, who did pioneering work in statistical physics, contributed greatly to the development of the atomic theory of gases which he put in its modern quantitative form. His work in 1872 led to fundamental insights into the microscopic explanation of irreversibility. He also laid the foundation for statistical mechanics and introduced the basic relation $S = k \ln \Omega$ connecting the entropy with the number of accessible states. Boltzmann's work came under heavy attack by a whole school of thought represented by such men as Ernst Mach (1838–1916) and Wilhelm Ostwald (1853–1932), who argued that physical theory should deal only with macroscopically observable quantities and should reject purely hypothetical concepts such as atoms. Discouraged, Boltzmann wrote in 1898, "I am conscious of being only an individual struggling weakly against the stream of time." Increasingly subject to depressions, Boltzmann committed suicide in 1906 only shortly before Perrin's experiment on Brownian motion (1908) and Millikan's oil-drop experiment (1909) provided very direct evidence for the discrete atomic structure of matter. (Photograph by courtesy of Professor W. Thirring, University of Vienna.)



J. Willard Gibbs

Fig. 4.9 Josiah Willard Gibbs (1839–1903). The first American theoretical physicist of note, he was born and died in New Haven where he spent all his working life as professor at Yale University. In the 1870s he made fundamental contributions to *thermodynamics*, couching its purely macroscopic reasoning in powerful analytical form and using it to treat many important problems of physics and chemistry. Around 1900 he developed a very general formulation of statistical mechanics in terms of the concept of ensembles. Despite the modifications brought about by quantum mechanics, the basic framework of his formulation still stands and is essentially the one upon which we have built our systematic discussion beginning with Chap. 3. The name *canonical ensemble* originated with Gibbs. (Photograph by courtesy of the Beinecke Rare Book and Manuscript Library, Yale University.)

4.6 Paramagnetism

The canonical distribution can be used to discuss a large number of situations of great physical interest. As a first application we shall investigate the magnetic properties of a substance which contains N_0 magnetic atoms per unit volume and which is placed in an external magnetic field B . We consider the particularly simple case where each magnetic atom has spin $\frac{1}{2}$ and an associated magnetic moment μ_0 . In a quantum-mechanical description the magnetic moment of each atom can then point either "up" (i.e., parallel to the external field) or "down" (i.e., antiparallel to the field). The substance is said to be *paramagnetic* since its magnetic properties are due to the orientation of individual magnetic moments. Suppose that the substance is at an absolute temperature T . What then is $\bar{\mu}$, the mean component of the magnetic moment of one of its atoms along the direction of the magnetic field B ?

We assume that each magnetic atom interacts only weakly with all the other atoms of the substance. In particular, we assume that the magnetic atoms are sufficiently far separated from each other so that we can neglect the magnetic field produced at the position of one magnetic atom by a neighboring magnetic atom. It is then permissible to focus attention on a single magnetic atom as the small system under consideration and to regard all the other atoms of the substance as constituting a heat reservoir at the absolute temperature T of interest.†

Each atom can be in two possible states: the state (+) where its magnetic moment points up, and the state (−) where its magnetic moment points down. Let us discuss these states in turn.

In the state (+), the atomic magnetic moment is parallel to the field so that $\mu = \mu_0$. The corresponding magnetic energy of the atom is then $\epsilon_+ = -\mu_0 B$. The canonical distribution (49) thus yields for the probability P_+ of finding the atom in this state the result

$$P_+ = Ce^{-\beta\epsilon_+} = Ce^{\beta\mu_0 B} \quad (54)$$

where C is a constant of proportionality and $\beta = (kT)^{-1}$. This is the state of lower energy and thus is the state in which the atom is more likely to be found.

† This assumes that it is possible to identify a single atom unambiguously, an assumption which is justified if the atoms are localized at definite lattice sites of a solid or if they form a dilute gas where the atoms are widely separated. In a sufficiently concentrated gas of identical atoms the assumption breaks down since the atoms are then indistinguishable in a quantum-mechanical description. It would then be necessary to adopt a point of view (which is always permissible, although more complicated) which considers the *entire* gas of atoms as a small macroscopic system in contact with some heat reservoir.

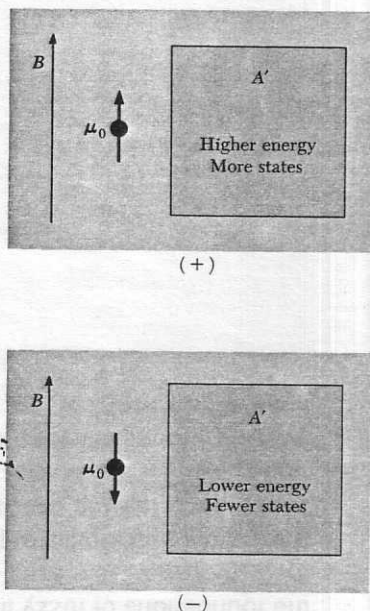


Fig. 4.10 An atom with spin $\frac{1}{2}$ shown in thermal contact with a heat reservoir A' . When the magnetic moment of the atom points up, its energy is less by an amount $2\mu_0 B$ than when it points down; accordingly the energy of the reservoir is greater by an amount $2\mu_0 B$ so that the reservoir can be in many more states. Hence the situation in which the moment points up occurs with greater probability than that in which it points down.

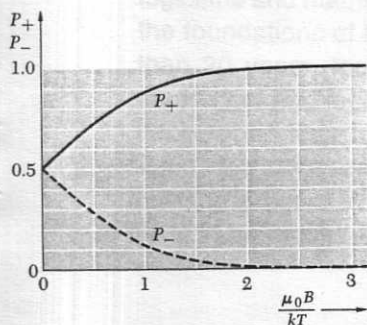


Fig. 4.11 Graph showing the probability P_+ that a magnetic moment μ_0 points parallel (and the probability P_- that it points antiparallel) to an applied magnetic field B when the absolute temperature is T .

In the state $(-)$ the atomic magnetic moment is antiparallel to the field so that $\mu = -\mu_0$. The corresponding energy of the atom is then $\epsilon_- = +\mu_0 B$. The probability P_- of finding the atom in this state is thus

$$P_- = C e^{-\beta \epsilon_-} = C e^{-\beta \mu_0 B}. \quad (55)$$

This is the state of higher energy and thus is the state in which the atom is less likely to be found.

The constant C is immediately determined by the normalization requirement that the probability of finding the atom in either one of these states must be unity. Thus

$$P_+ + P_- = C(e^{\beta \mu_0 B} + e^{-\beta \mu_0 B}) = 1$$

or

$$C = \frac{1}{e^{\beta \mu_0 B} + e^{-\beta \mu_0 B}}. \quad (56)$$

Since the atom is more likely to be in the state $(+)$ where its magnetic moment is parallel to the field B , the mean magnetic moment $\bar{\mu}$ must point in the direction of the field B . By virtue of (54) and (55), the significant parameter characterizing the orientation of the magnetic moment is the quantity

$$w \equiv \beta \mu_0 B = \frac{\mu_0 B}{kT} \quad (57)$$

which measures the ratio of the magnetic energy $\mu_0 B$ to the characteristic thermal energy kT . It is apparent that if T is very large (i.e., if $w \ll 1$), the probability that the magnetic moment is parallel to the field is almost the same as that it is antiparallel. In this case the magnetic moment is almost completely randomly oriented so that $\bar{\mu} \approx 0$. On the other hand, if T is very small (i.e., if $w \gg 1$), it is much more probable that the magnetic moment is parallel to the field than antiparallel to it. In this case $\bar{\mu} \approx \mu_0$.

All these qualitative conclusions can readily be made quantitative by actually calculating the mean value $\bar{\mu}$. Thus we find

$$\bar{\mu} \equiv P_+(\mu_0) + P_-(-\mu_0) = \mu_0 \frac{e^{\beta \mu_0 B} - e^{-\beta \mu_0 B}}{e^{\beta \mu_0 B} + e^{-\beta \mu_0 B}}. \quad (58)$$

Alternatively, this result can be written in the form

$$\bar{\mu} = \mu_0 \tanh\left(\frac{\mu_0 B}{kT}\right) \quad (59)$$

where we have used the definition of the hyperbolic tangent

$$\tanh w \equiv \frac{e^w - e^{-w}}{e^w + e^{-w}}. \quad (60)$$

- .. The mean magnetic moment per unit volume of the substance (i.e., its *magnetization*) points then in the direction of the magnetic field. Its magnitude $\bar{M}_0$ is simply given by

$$\bar{M}_0 = N_0 \bar{\mu} \quad (61)$$

if there are N_0 magnetic atoms per unit volume.

- * We can easily verify that $\bar{\mu}$ exhibits the behavior discussed previously in qualitative terms. If $w \ll 1$, then $e^w = 1 + w + \dots$ and $e^{-w} = 1 - w + \dots$. Hence,

$$\text{for } w \ll 1, \quad \tanh w = \frac{(1 + w + \dots) - (1 - w + \dots)}{2} = w.$$

On the other hand, if $w \gg 1$, then $e^w \gg e^{-w}$. Hence,

$$\text{for } w \gg 1, \quad \tanh w = 1.$$

The relation (59) predicts thus the following limiting behavior:

$$\text{for } \mu_0 B \ll kT, \quad \bar{\mu} = \mu_0 \left(\frac{\mu_0 B}{kT} \right) = \frac{\mu_0^2 B}{kT}; \quad (62)$$

$$\text{for } \mu_0 B \gg kT, \quad \bar{\mu} = \mu_0. \quad (63)$$

- .. When $\mu_0 B \ll kT$, the value of $\bar{\mu}$ is rather small. By (62), $\bar{\mu}$ is then less than its maximum possible value μ_0 by the ratio $(\mu_0 B/kT)$. Note that $\bar{\mu}$ in this limit is simply proportional to the magnetic field B and inversely proportional to the absolute temperature T . Using (61) and (62), the magnetization then becomes,

$$\text{for } \mu_0 B \ll kT, \quad \bar{M}_0 = N_0 \bar{\mu} = \frac{N_0 \mu_0^2 B}{kT} \equiv \chi B \quad (64)$$

where χ is a constant of proportionality independent of B . This parameter χ is called the *magnetic susceptibility* of the substance.†

† The magnetic susceptibility is conventionally defined in terms of the magnetic field H so that $\chi \equiv \bar{M}_0/H$. But since the concentration N_0 of magnetic atoms is supposed to be small, $H = B$ to excellent approximation.

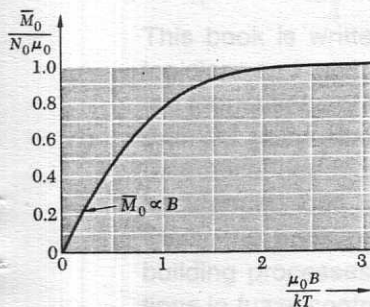


Fig. 4.12 Dependence of the magnetization $\bar{M}_0$ on the magnetic field B and temperature T in the case of weakly interacting magnetic atoms of spin $\frac{1}{2}$ and magnetic moment μ_0 .

Equation (64) thus provides the following explicit expression for χ in terms of microscopic quantities:

$$\chi = \frac{N_0 \mu_0^2}{kT}. \quad (65)$$

The fact that χ is inversely proportional to the absolute temperature is known as *Curie's law*.

When $\mu_0 B \gg kT$, the mean magnetic moment $\bar{\mu}$ attains its maximum possible value μ_0 . Correspondingly the magnetization becomes

$$\text{for } \mu_0 B \gg kT, \quad \bar{M}_0 = N_0 \mu_0 \quad (66)$$

which is its maximum possible (or *saturation*) value and is thus independent of B or T . The complete dependence of the magnetization $\bar{M}_0$ on the absolute temperature T and magnetic field B is shown in Fig. 4.12.

4.7 Mean Energy of an Ideal Gas

Consider a gas of N identical molecules, each of mass m , confined within a box of edge lengths L_x , L_y , and L_z . Suppose that the gas is sufficiently dilute, i.e., that the number N of molecules in the given volume $V = L_x L_y L_z$ is sufficiently small so that the mean separation between the molecules is correspondingly large. The following two simplifying conditions are then satisfied:

- (i) The mean potential energy of mutual interaction between the molecules is very small compared to their mean kinetic energy. (The gas is then said to be *ideal*.)
- (ii) It is permissible to focus attention on any one particular molecule as an entity which can be identified despite the essential indistinguishability of the molecules. (The gas is then said to be *nondegenerate*.)†

We shall assume that the gas is sufficiently dilute so that both of these conditions are satisfied.‡

† This last conclusion is based on the fact that the mean separation between the molecules is large compared to the typical de Broglie wavelength of a molecule. If this is not the case, quantum-mechanical limitations make it impossible to focus attention unambiguously on a particular molecule and a rigorous quantum-mechanical treatment of a system of indistinguishable particles becomes essential. (The gas is then said to be *degenerate* and is described by either the so-called Bose-Einstein or Fermi-Dirac statistics.)

‡ Condition (ii) is satisfied by practically all ordinary gases; the extent of its validity will be examined more quantitatively at the end of Sec. 6.3. When a gas is made less dilute, condition (i) usually breaks down long before condition (ii) is violated. If the interaction between molecules is very small, however, a gas may satisfy condition (i) so that it is ideal, but violate condition (ii).

Suppose that the gas is in equilibrium at an absolute temperature T . By virtue of condition (ii), we can focus attention on a particular molecule of the gas and regard it as a small system in thermal contact with a heat reservoir (at temperature T) consisting of all the other molecules of the gas. The probability of finding the molecule in any one of its quantum states r , where its energy is ϵ_r , is then simply given by the canonical distribution (49) or (51); i.e.,

$$P_r = \frac{e^{-\beta\epsilon_r}}{\sum_r e^{-\beta\epsilon_r}}, \quad \text{where } \beta \equiv \frac{1}{kT}. \quad (67)$$

In calculating the energy ϵ_r of the molecule, the condition (i) allows us to neglect any energy of interaction of the molecule with other molecules.

Consider, for instance, the particularly simple case of a *monatomic* gas [such as helium (He) or argon (Ar)] in which each molecule consists of a single atom only. The energy of such a molecule is then simply its kinetic energy. Each possible quantum state r of the molecule is accordingly labeled by some particular values of the three quantum numbers $\{n_x, n_y, n_z\}$ and has an energy given by Eq. (3.15). Thus

$$\epsilon_r = \frac{\pi^2 \hbar^2}{2m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right). \quad (68)$$

The probability of finding the molecule in any such state is then given by (67).

On the other hand, consider the case of a *polyatomic* gas [such as oxygen (O_2), nitrogen (N_2), or methane (CH_4)] in which each molecule consists of two or more atoms. Then the energy ϵ of each molecule is given by

$$\epsilon = \epsilon^{(k)} + \epsilon^{(i)}. \quad (69)$$

Here $\epsilon^{(k)}$ is the kinetic energy of translational motion of the center of mass of the molecule, while $\epsilon^{(i)}$ is the intramolecular energy associated with the rotation and vibration of the atoms with respect to the center of mass. Since the center of mass moves like a simple particle having the mass of the molecule, the state of translational motion of the molecule is again specified by the set of three quantum numbers $\{n_x, n_y, n_z\}$ and its translational kinetic energy $\epsilon^{(k)}$ is again given by (68). The state of intramolecular motion is specified by one or more other quantum numbers (denote them collectively simply by n_i) describing the state of rotational and vibrational motion of the atoms in the molecule; the energy $\epsilon^{(i)}$ depends then on n_i . A particular state r of the molecule is

additional degrees of freedom

then specified by specific values of the quantum numbers $\{n_x, n_y, n_z; n_i\}$ and has a corresponding energy ϵ_r given by

$$\epsilon_r = \epsilon^{(k)}(n_x, n_y, n_z) + \epsilon^{(i)}(n_i). \quad (70)$$

Note that the translational motion of the molecule is affected by the presence of the confining walls of the container; hence $\epsilon^{(k)}$ depends on the dimensions L_x, L_y, L_z of the container, as seen explicitly in (68).

On the other hand, the intramolecular motion of the atoms with respect to the molecular center of mass does *not* involve the dimensions of the container; hence $\epsilon^{(i)}$ is independent of the dimensions of the container, i.e.,

$$\epsilon^{(i)} \text{ is independent of } L_x, L_y, L_z. \quad (71)$$

Calculation of the mean energy

If a molecule is found with probability P_r in a state r of energy ϵ_r , then its mean energy is simply given by

$$\bar{\epsilon} = \sum_r P_r \epsilon_r = \frac{\sum_r e^{-\beta \epsilon_r} \epsilon_r}{\sum_r e^{-\beta \epsilon_r}} \quad (72)$$

where we have used (67) and where the sums are over all possible states r of the molecule. The relation (72) can be considerably simplified by noting that the sum in the numerator can be readily expressed in terms of the sum in the denominator. Thus we can write

$$\sum_r e^{-\beta \epsilon_r} \epsilon_r = - \sum_r \frac{\partial}{\partial \beta} (e^{-\beta \epsilon_r}) = - \frac{\partial}{\partial \beta} \left(\sum_r e^{-\beta \epsilon_r} \right)$$

where we have used the fact that the derivative of a sum of terms is equal to the sum of the derivatives of these terms. Introducing for the denominator in (72) the convenient abbreviation

$$Z \equiv \sum_r e^{-\beta \epsilon_r}, \quad (73)$$

the relation (72) becomes then

$$\bar{\epsilon} = \frac{- \frac{\partial Z}{\partial \beta}}{Z} = - \frac{1}{Z} \frac{\partial Z}{\partial \beta}$$

or

$$\bar{\epsilon} = - \frac{\partial \ln Z}{\partial \beta}. \quad (74)$$

The calculation of the mean energy $\bar{\epsilon}$ thus requires only the evaluation of the single sum Z of (73). (The sum Z over all states of the molecule is called the *partition function* of the molecule.)

In the case of a monatomic gas, the energy levels are given by (68) and the sum Z of (73) becomes accordingly†

$$Z = \sum_{n_x} \sum_{n_y} \sum_{n_z} \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right) \right] \quad (75)$$

where the triple sum is over all possible values of n_x , n_y , and n_z [each ranging over all integral values from 1 to ∞ in accordance with (3.14)]. But the exponential function factors immediately into a simple product of exponentials, i.e.,

$$\begin{aligned} & \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right) \right] \\ &= \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \frac{n_x^2}{L_x^2} \right] \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \frac{n_y^2}{L_y^2} \right] \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \frac{n_z^2}{L_z^2} \right]. \end{aligned}$$

Here n_x appears only in the first factor, n_y only in the second, and n_z only in the third. Hence the sum (75) breaks up into a simple product, i.e.,

$$Z = Z_x Z_y Z_z \quad (76)$$

$$\text{where} \quad Z_x \equiv \sum_{n_x=1}^{\infty} \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \frac{n_x^2}{L_x^2} \right], \quad (77a)$$

$$Z_y \equiv \sum_{n_y=1}^{\infty} \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \frac{n_y^2}{L_y^2} \right], \quad (77b)$$

$$Z_z \equiv \sum_{n_z=1}^{\infty} \exp \left[-\frac{\beta \pi^2 \hbar^2}{2m} \frac{n_z^2}{L_z^2} \right]. \quad (77c)$$

It only remains to evaluate a typical sum such as Z_x . This is easily done if we note that, for any container where L_x is of macroscopic size, the coefficient of n_x^2 in (77a) is very small unless β is extremely large (i.e., unless T is extremely small). Since successive terms in the sum differ thus only very slightly in magnitude, it is an excellent approximation to replace this sum by an integral. By regarding a term in the

† To prevent the long expression in the exponent from overwhelming the poor letter e , we here use the standard alternative notation $\exp u \equiv e^u$ to denote the exponential function.

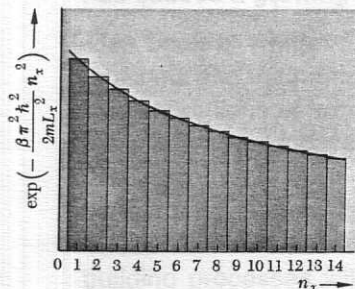


Fig. 4.13 Schematic diagram illustrating the replacement of a sum over integral values of n_x (the sum being equal to the area beneath the rectangles) by an integral over continuous values of n_x (the integral being equal to the area beneath the curve).

sum as a function of n_x (considered as a continuous variable defined also for nonintegral values), the sum Z_x thus becomes

$$Z_x = \int_{1/2}^{\infty} \exp \left[-\frac{\beta \pi^2 \hbar^2 n_x^2}{2m L_x^2} \right] dn_x \\ = \left(\frac{2m}{\beta} \right)^{1/2} \left(\frac{L_x}{\pi \hbar} \right) \int_0^{\infty} \exp [-u^2] du, \quad (78)$$

where
$$u \equiv \left(\frac{\beta}{2m} \right)^{1/2} \left(\frac{\pi \hbar}{L_x} \right) n_x \quad (79)$$

or
$$n_x = \left(\frac{2m}{\beta} \right)^{1/2} \left(\frac{L_x}{\pi \hbar} \right) u.$$

The lower limit in the last integral of (78) has been put equal to zero without noticeable error since the coefficient of n_x in (79) is extremely small. The last definite integral in (78) is merely equal to some constant so that (78) is of the form

$$Z_x = b \frac{L_x}{\beta^{1/2}} \quad (80)$$

where b is some constant involving the mass of the molecule.† The corresponding expressions for Z_y and Z_z are, of course, analogous to (80). Hence (76) yields the result

$$Z = \left(b \frac{L_x}{\beta^{1/2}} \right) \left(b \frac{L_y}{\beta^{1/2}} \right) \left(b \frac{L_z}{\beta^{1/2}} \right)$$

or
$$Z = b^3 \frac{V}{\beta^{3/2}}, \quad \text{where } V \equiv L_x L_y L_z \quad (81)$$

is the volume of the box. Correspondingly, we obtain

$$\ln Z = \ln V - \frac{3}{2} \ln \beta + 3 \ln b. \quad (82)$$

Our calculation is now essentially completed. Indeed, (74) yields for the mean energy $\bar{\epsilon}$ of a molecule the result

$$\bar{\epsilon} = -\frac{\partial \ln Z}{\partial \beta} = -\left(-\frac{3}{2} \frac{1}{\beta} \right) = \frac{3}{2} \left(\frac{1}{\beta} \right).$$

We have thus arrived at the important conclusion that

$$\bar{\epsilon} = \frac{3}{2} kT. \quad (83)$$

† Although this will be of no particular importance to us, we note that the last integral in (78) has, by (M.21), the value $\sqrt{\pi}/2$; thus $b = (m/2\pi)^{1/2} \hbar^{-1}$.

- * The mean kinetic energy of a molecule is thus independent of the size of the container and simply proportional to the absolute temperature T of the gas.

If the molecules of the gas are not monatomic, the additive expression (69) yields for the mean energy of a molecule the result

$$\bar{\epsilon} = \bar{\epsilon}^{(k)} + \bar{\epsilon}^{(i)} = \frac{3}{2} kT + \bar{\epsilon}^{(i)}(T) \quad (84)$$

- since the mean kinetic energy of translation $\bar{\epsilon}^{(k)}$ of the center of mass is again given by (83). The mean intramolecular energy $\bar{\epsilon}^{(i)}$ is, by (71), independent of the dimensions of the container and thus can only be a function of the absolute temperature T .

Since the gas is ideal so that the interaction between the molecules is negligible, the total mean energy E of the gas is simply equal to the sum of the mean energies of the N individual molecules. Thus we have

$$E = N\bar{\epsilon}. \quad (85)$$

- Even in the most general case, the mean energy of an ideal gas is independent of the dimensions of the container and a function of the temperature only; i.e.,

for an ideal gas,

$$E = E(T) \quad (86)$$

independent of the dimensions of the container.

- † This result is physically plausible. The translational kinetic energy and intramolecular energy of a molecule do not depend on the separations between the molecules. Hence a change of the container dimensions (at the fixed temperature T) leaves these energies unchanged so that E remains unaffected. If the gas were not ideal, this conclusion would no longer be true. Indeed, if the gas is sufficiently dense, the mean separation between molecules is small enough so that their mean potential energy of mutual interaction is appreciable. A change of the container dimensions (at the fixed temperature T) then results in a change in the mean separation of the molecules; correspondingly, it affects their mean intermolecular *potential* energy which is included in the total mean energy E of the gas.

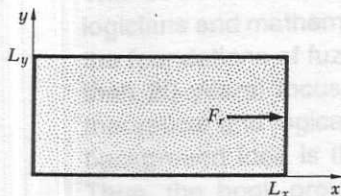


Fig. 4.14 A box containing an ideal gas. A molecule in a given state r exerts on the right wall of the box a force F_r in the x direction.

4.8 Mean Pressure of an Ideal Gas

It is experimentally quite easy to measure the mean pressure (i.e., the mean force per unit area) exerted by a gas on the walls of the container within which it is confined. Hence it is of particular interest to calculate the mean pressure exerted by an ideal gas. Let us denote by F the force in the x direction exerted by a single molecule on the right wall (i.e., the wall $x = L_x$) of the box containing the gas. Let us denote by F_r the value of this force when the molecule is in a particular quantum state r where its energy is ϵ_r . The force F_r can be readily related to the energy ϵ_r . Indeed, suppose that the right wall of the box was displaced very slowly to the right by an amount dL_x . In this process the molecule would do on the wall an amount of work $F_r dL_x$ which must be equal to the decrease $-d\epsilon_r$ of the energy of the molecule. Hence

$$F_r dL_x = -d\epsilon_r$$

or

$$F_r = -\frac{\partial \epsilon_r}{\partial L_x} \quad (87)$$

Here we have written a partial derivative to indicate that the other dimensions L_y and L_z were supposed to remain constant in the argument leading to (87).

The mean force $\bar{F}$ exerted by a molecule on the wall is then given by averaging the force F_r over all possible states r of the molecule. Thus

$$\bar{F} = \frac{\sum_r P_r F_r}{\sum_r P_r} = \frac{\sum_r e^{-\beta \epsilon_r} \left(-\frac{\partial \epsilon_r}{\partial L_x} \right)}{\sum_r e^{-\beta \epsilon_r}} \quad (88)$$

where we have used the expression (67) for the probability P_r of finding the molecule in any state r . The relation (88) can be simplified since the sum in the numerator can again be expressed in terms of the sum in the denominator. Thus the numerator can be written

$$\begin{aligned} -\sum_r e^{-\beta \epsilon_r} \frac{\partial \epsilon_r}{\partial L_x} &= -\sum_r \left(-\frac{1}{\beta} \right) \frac{\partial}{\partial L_x} (e^{-\beta \epsilon_r}) \\ &= \frac{1}{\beta} \frac{\partial}{\partial L_x} \left(\sum_r e^{-\beta \epsilon_r} \right). \end{aligned}$$

Using again the abbreviation Z of (73), the expression (88) becomes

$$F = \frac{\frac{1}{\beta} \frac{\partial Z}{\partial L_x}}{Z} = \frac{1}{\beta} \frac{1}{Z} \frac{\partial Z}{\partial L_x}$$

or

$$\boxed{F = \frac{1}{\beta} \frac{\partial \ln Z}{\partial L_x}} \quad (89)$$

This general relation can now be applied to the result (82) already obtained for $\ln Z$ in the case of a monatomic molecule. Remembering that $V = L_x L_y L_z$, the partial differentiation yields immediately

$$F = \frac{1}{\beta} \frac{\partial \ln Z}{\partial L_x} = \frac{1}{\beta} \frac{\partial \ln V}{\partial L_x} = \frac{1}{\beta L_x}$$

or

$$\boxed{\bar{F} = \frac{kT}{L_x}} \quad (90)$$

† If the molecule is not monatomic, the expression (87) for the force F_r becomes by (70)

$$F_r = - \frac{\partial}{\partial L_x} [\epsilon_r^{(k)} + \epsilon_r^{(i)}] = - \frac{\partial \epsilon_r^{(k)}}{\partial L_x}.$$

Here we have used the fact, noted in (71), that the intramolecular energy $\epsilon^{(i)}$ does not depend on the dimension L_x of the box. Hence the force F_r is calculable from the center-of-mass translational energy alone. The preceding computation, which was based on this translational energy, remains therefore equally valid for a polyatomic molecule. The expression (90) for $\bar{F}$ is accordingly a completely general result.

† Since the gas is ideal, the molecules move about without influencing each other appreciably. Hence the *total* mean normal force (i.e., force in the x direction) exerted on the right wall by all the molecules of the gas is simply obtained by multiplying {the mean force $\bar{F}$ exerted by one molecule} by {the total number N of molecules in the gas}. Dividing this result by the area $L_y L_z$ of the wall gives then the mean pressure $\bar{p}$ exerted by the gas on this wall. The relation (90) thus leads to the result

$$\bar{p} = \frac{N\bar{F}}{L_y L_z} = \frac{N}{L_y L_z} \frac{kT}{L_x} = \frac{N}{V} kT.$$

Hence

$$\bar{p}V = NkT \quad (91)$$

or

$$\bar{p} = nkT \quad (92)$$

where $V = L_x L_y L_z$ is the volume of the container and where $n \equiv N/V$ is the number of molecules per unit volume. Note that (92) makes no specific reference to the particular wall which was considered in the calculation. The calculation would thus yield properly the same result for the mean pressure $\bar{p}$ exerted on *any* wall.† ‡

Discussion

The important relations (91) and (92) can be expressed in another equivalent form. Thus the total number N of molecules is ordinarily deduced from macroscopic measurements of the number ν of moles of gas present in the container. Since the number of molecules per mole is, by definition, equal to Avogadro's number N_a , it follows that $N = \nu N_a$. Hence (91) can also be written in the form

$$\bar{p}V = \nu RT \quad (93)$$

where we have introduced a new constant R , called the *gas constant*, defined by the relation

$$R \equiv N_a k. \quad (94)$$

The equation relating the pressure, the volume, and the absolute temperature of a substance in equilibrium is called the *equation of state* of that substance. Thus the equations (91) through (93) are dif-

† Indeed, an elementary (and completely macroscopic) analysis of forces in a fluid in mechanical equilibrium shows that the pressure on any element of area must be the same anywhere in the fluid (if gravity is neglected) and must be independent of the spatial orientation of this element of area.

‡ Note concerning Secs. 4.7 and 4.8:

Although our calculations of the mean energy and pressure have been carried through for a gas confined within a container in the simple shape of a rectangular parallelepiped, all our results are actually quite general and independent of the shape of the container.

The physical reason is the following. At any ordinary temperature the momentum of a molecule is so large that its de Broglie wavelength is negligibly small compared to the dimensions of any macroscopic container. Practically every region within the container is thus many wavelengths away from the container walls. Accordingly, the nature of the possible wave functions which can exist in such a region is utterly insensitive to the detailed boundary conditions imposed at the walls or to details involving the precise shape of these walls.

ferent forms of the equation of state for an ideal gas. This equation of state, which we have derived on the basis of our theory, makes several important predictions:

(i) If a given amount of gas, sufficiently dilute to be ideal, is maintained at a constant temperature, it follows by (91) that

$$\bar{p}V = \text{constant},$$

i.e., that its pressure should be inversely proportional to its volume. This result was discovered experimentally by Boyle in 1662 (long before the advent of the atomic theory of matter) and is accordingly called *Boyle's law*.

(ii) If a given amount of gas, sufficiently dilute to be ideal, is maintained at constant volume, its mean pressure should be proportional to its absolute temperature. This result can be conveniently exploited, as will be shown in the next chapter, as a method for measuring the absolute temperature.

(iii) The equation of state (91) depends only on the number of molecules, but makes no reference to the nature of these molecules. Hence the equation of state should be the same for *any* gas (whether it be He, H₂, N₂, O₂, CH₄, etc.), as long as it is made sufficiently dilute to be ideal. This prediction can be checked experimentally and is well verified.

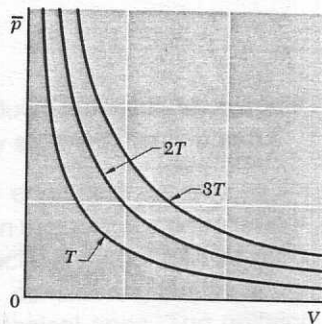


Fig. 4.15 Dependence of the mean pressure $\bar{p}$ of an ideal gas on its volume V at the absolute temperatures T , $2T$, and $3T$.

Summary of Definitions

absolute temperature The absolute temperature T of a macroscopic system [or its related parameter $\beta = (kT)^{-1}$] is defined by

$$\frac{1}{kT} \equiv \beta \equiv \frac{\partial \ln \Omega}{\partial E}$$

where $\Omega(E)$ is the number of states accessible to the system in a small energy interval between E and $E + \delta E$, and where k is a conventionally chosen constant called *Boltzmann's constant*.

entropy The entropy S of a system is defined in terms of the number Ω of states accessible to it by the relation

$$S \equiv k \ln \Omega.$$

The entropy thus provides a logarithmic measure of the degree of randomness of the system.

thermometer A relatively small macroscopic system arranged so that only one of its macroscopic parameters changes when the system gains or loses energy as a result of thermal interaction.

thermometric parameter The variable macroscopic parameter of a thermometer.

temperature of a system with respect to a given thermometer The value of the particular thermometric parameter of this thermometer when the latter is in thermal equilibrium with the system.

heat reservoir A macroscopic system which is sufficiently large with respect to a set of other systems under consideration so that its temperature remains essentially unchanged in any thermal interaction with those systems.

Boltzmann factor The factor $e^{-\beta E}$ where β is related to the absolute temperature T by $\beta \equiv (kT)^{-1}$ and where E denotes an energy.

canonical distribution The probability distribution according to which the probability P_r of finding a system in a state r of energy E_r is given by

$$P_r \propto e^{-\beta E_r}$$

where $\beta = (kT)^{-1}$ is the absolute temperature parameter of the heat reservoir with which the system is in equilibrium.

ideal gas A gas in which the energy of interaction between the molecules is almost negligible compared to their kinetic energy.

nondegenerate gas A gas sufficiently dilute so that the mean separation between molecules is large compared to the mean de Broglie wavelength of a molecule.

equation of state The relation connecting the volume, the mean pressure, and the absolute temperature of a given macroscopic system.

Important Relations

Definition of absolute temperature:

$$\frac{1}{kT} \equiv \beta \equiv \frac{\partial \ln \Omega}{\partial E}. \quad (\text{i})$$

Definition of entropy:

$$S \equiv k \ln \Omega. \quad (\text{ii})$$

Entropy increase of a system at absolute temperature T when it absorbs a small amount of heat dQ :

$$dS = \frac{dQ}{T}. \quad (\text{iii})$$

Canonical distribution for a system in thermal equilibrium with a heat reservoir at absolute temperature T :

$$P_r \propto e^{-\beta E_r}. \quad (\text{iv})$$

Equation of state of an ideal nondegenerate gas:

$$\bar{p} = nkT. \quad (\text{v})$$

Suggestions for Supplementary Reading

The following books provide alternative derivations of some of the results of this chapter:

- C. W. Sherwin, *Basic Concepts of Physics*, secs. 7.3–7.5 (Holt, Rinehart and Winston, Inc., New York, 1961).
- G. S. Rushbrooke, *Introduction to Statistical Mechanics*, chaps. 2 and 3 (Oxford University Press, Oxford, 1949).
- F. C. Andrews, *Equilibrium Statistical Mechanics*, secs. 6–8 (John Wiley & Sons, Inc., New York, 1963).

Historical and biographical accounts:

- H. Thirring, "Ludwig Boltzmann," *J. of Chemical Education*, p. 298 (June 1952).
- E. Broda, Ludwig Boltzmann; Mensch, Physiker, Philosoph (Franz Deuticke, Vienna, 1955). In German.
- L. Boltzmann, *Lectures on Gas Theory*, translated by S. G. Brush (University of California Press, Berkeley, 1964). The introduction by Brush describes briefly the historical development of the description of matter in terms of atomic concepts.
- B. A. Leeburger, Josiah Willard Gibbs, *American Theoretical Physicist* (Franklin Watts, Inc., New York, 1963).
- L. P. Wheeler, Josiah Willard Gibbs, *the History of a Great Mind* (Yale University Press, New Haven, 1951; paperback ed., 1962).
- M. Rukeyser, *Willard Gibbs* (Doubleday & Company, Inc., Garden City, N.Y., 1942).

Problems

4.1 Example of a peculiar thermometer

The density of alcohol, like that of most substances, is found to decrease with increasing absolute temperature. The properties of water are, however, somewhat unusual. As the absolute temperature is increased above the temperature at which water melts (i.e., turns from ice into liquid water), its density at first increases and then, after going through a maximum, decreases.

Suppose that an ordinary thermometer, of the type consisting of a liquid column within a glass tube, is filled with colored water instead of the colored alcohol which is commonly used. The temperature θ indicated by this type of thermometer is, as usual, taken to be the length of the liquid column. Imagine that this thermometer, when placed in contact with two systems A and B , indicates that they have respective temperatures θ_A and θ_B .

(a) Suppose that the temperature θ_A of system A is larger (higher) than the temperature θ_B of system B . Can one necessarily conclude that heat will flow from system A to system B when these two systems are placed in thermal contact with each other?

(b) Suppose that the temperatures θ_A and θ_B of the two systems are found to be equal. Can one necessarily conclude that no heat will flow from system A to system B when these two systems are placed in thermal contact with each other?

4.2 Value of kT at room temperature

One mole of any gas at room temperature and atmospheric pressure (10^6 dynes/cm²) is found experimentally to occupy a volume of approximately 24 liters (i.e., 24×10^3 cm³). Use this result to estimate the value of kT at room temperature. Express your answer in ergs and also in electron volts. (1 electron volt = 1.60×10^{-12} erg.)

4.3 Actual variation of the number of states with energy

Consider any macroscopic system at room temperature.

(a) Use the definition of absolute temperature to find the percentage increase in the number of states accessible to such a system when its energy is increased by 10^{-3} electron volt.

(b) Suppose that such a system absorbs a single photon of visible light (having a wavelength of 5×10^{-5} cm). By what factor does the number of states accessible to the system increase as a result?

4.4 Achievement of atomic spin polarization

Consider a substance which contains magnetic atoms having a spin $\frac{1}{2}$ and a magnetic moment μ_0 . Since this moment is due to an unpaired electron, it is of the order of a Bohr magneton, i.e., $\mu_0 \approx 10^{-20}$ erg/gauss. In order to do scattering experiments from atoms whose spins are preferentially polarized in a given direction, one can apply a large magnetic field B and cool the substance

to a sufficiently low absolute temperature to achieve appreciable polarization.

The largest magnetic field which can conveniently be produced in the laboratory is about 50,000 gauss. Find the absolute temperature T which must be attained so that the number of atomic moments pointing parallel to the field is at least 3 times as large as the number of moments pointing in the opposite direction. Express your answer in terms of the ratio T/T_R where T_R is room temperature.

4.5 Proposed method for producing polarized proton targets

In research in nuclear physics and elementary particles, it is of great interest to do scattering experiments on targets consisting of protons whose spins are preferentially polarized in a given direction. Each proton has a spin $\frac{1}{2}$ and a magnetic moment $\mu_0 = 1.4 \times 10^{-23}$ erg/gauss. Suppose that one tries to apply the method of the preceding problem by taking a sample of paraffin (which contains many protons), applying a magnetic field of 50,000 gauss, and cooling the sample to some very low absolute temperature T . How low would this temperature have to be so that, after equilibrium has been reached, the number of proton moments pointing parallel to the field is at least 3 times as large as the number of proton moments pointing in the opposite direction? Express your answer again in terms of the ratio T/T_R where T_R is room temperature.

4.6 Nuclear magnetic resonance absorption

A sample of water is placed in an external magnetic field B . Each proton of the H_2O molecule has a nuclear spin $\frac{1}{2}$ and a small magnetic moment μ_0 . Since each proton can point either "up" or "down," it can be in one of two possible states of respective energies $\mp\mu_0 B$. Suppose that one applies a radio-frequency magnetic field of frequency ν which is such that it satisfies the resonance condition $h\nu = 2\mu_0 B$, where $2\mu_0 B$ is the energy difference between these two proton states and h is Planck's constant. Then the radiation field produces transitions between these two states, causing the proton to go from the "up" state to the "down" state, or vice versa, with equal probability. The net power absorbed by the protons from the radiation field is then proportional to the difference between the numbers of protons in the two states.

Assume that the protons always remain very close to equilibrium at the absolute temperature T of the water. How does the absorbed power depend on the temperature T ? Use the excellent approximation based on the fact that μ_0 is so small that $\mu_0 B \ll kT$.

4.7 Relative numbers of atoms in an atomic beam experiment

Precision measurements of the magnetic moment of the electron were instrumental in leading to our modern understanding of the quantum theory of the electromagnetic field. The first such precision experiment, performed by Kusch and Foley [*Physical Review* 74, 250 (1948)], was based on a comparison of the measured total magnetic moments of gallium (Ga) atoms in two different sets of states labeled (in standard spectroscopic notation) as the $^2P_{1/2}$ and $^2P_{3/2}$

states, respectively. The ${}^2P_{1/2}$ states are the states of lowest possible energy of the atom. (There are *two* such states with the same energy; they correspond merely to the two possible spatial orientations of the total angular momentum of the atom in this set of states.) The ${}^2P_{3/2}$ states have an energy higher than that of the ${}^2P_{1/2}$ states by an amount accurately known from spectroscopic measurements and equal to 0.102 electron volt. (There are *four* such states having the same energy; they correspond again merely to the four possible spatial orientations of the total angular momentum of the atom in this set of states.)

To perform the desired comparison, the number of atoms in the ${}^2P_{3/2}$ states must be comparable to the number of atoms in the ${}^2P_{1/2}$ states. The atoms can be produced by heating gallium in an oven to a high absolute temperature T . A small hole in the oven wall then permits a few of the atoms to escape into a surrounding vacuum where they form an *atomic beam* on which the actual measurements are made.

(a) Suppose that the absolute temperature T of the oven is $3T_R$, where T_R is room temperature. What is then the proportion of gallium atoms in the beam which are in the ${}^2P_{1/2}$ and ${}^2P_{3/2}$ states, respectively?

(b) The highest temperature which can be produced conveniently in such an oven is about $6T_R$. What then is the proportion of gallium atoms in the ${}^2P_{1/2}$ and ${}^2P_{3/2}$ states, respectively? Is this proportion adequate for a successful experiment?

4.8 Mean energy of a system with two discrete energy levels

A system consists of N weakly interacting particles, each of which can be in either one of two states with respective energies ϵ_1 and ϵ_2 , where $\epsilon_1 < \epsilon_2$.

(a) Without explicit calculation, make a qualitative plot of the mean energy E of the system as a function of its absolute temperature T . What is E in the limit of very low and very high temperatures? Roughly near what temperature does E change from its low to its high temperature limiting values?

(b) Find an explicit expression for the mean energy of this system. Verify that this expression exhibits the temperature dependence established qualitatively in (a).

4.9 Elastic properties of rubber

A strip of rubber maintained at an absolute temperature T is fastened at one end to a peg; from its other end hangs a weight w . Assume the following simple microscopic model for the rubber band: It consists of a linked polymer chain of N segments joined end to end; each segment has length a and can be oriented either parallel or antiparallel to the vertical direction. Find an expression for the resultant mean length L of the rubber band as a function of w . (Neglect the kinetic energies or weights of the segments themselves, or any interaction between the segments.)

4.10 Polarization due to impurity atoms in a solid

The following represents a simple two-dimensional model of a situation of actual physical interest. A solid at absolute temperature T contains N_0 negatively charged impurity ions per unit volume, these ions replacing some of the ordinary atoms of the solid. The solid as a whole is, of course, electrically neutral. This is so because each negative ion with charge $-e$ has in its vicinity one positive ion with charge $+e$. The positive ion is small and thus free to move between lattice sites. In the absence of an external electric field it will, therefore, be found with equal probability in any one of the four equidistant sites surrounding the stationary negative ion. (See Fig. 4.16; the lattice spacing is a .)

If a small electrical field $\mathcal{E}$ is applied along the x direction, calculate the electric polarization, i.e., the mean electric dipole moment per unit volume along the x direction.

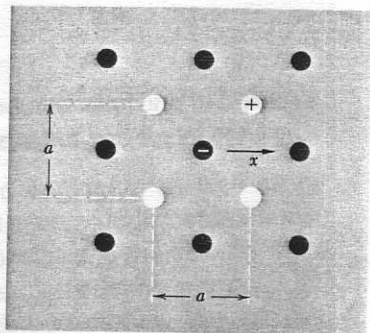


Fig. 4.16 Impurity atoms in the crystal lattice of a solid.

4.11 Minimum property of the "free energy" for a system in contact with a heat reservoir

When two systems A and A' are placed in thermal contact, their total entropy tends to increase in accordance with the relation (20), i.e.,

$$\Delta S + \Delta S' \geq 0. \quad (i)$$

The equilibrium situation ultimately attained after system A has absorbed some heat $Q = \Delta E$ thus corresponds to that where the total entropy $S + S'$ of the combined isolated system is maximum.

Suppose now that A is small compared to A' so that A' acts as a heat reservoir at some constant absolute temperature T' . The entropy change $\Delta S'$ of A' can then be expressed very simply in terms of ΔE and T' . Show that (i) implies in this case that the quantity $F \equiv E - T'S$ tends to decrease and becomes a minimum in the equilibrium situation. (The function F is called the Helmholtz free energy of the system A at the constant temperature T' .)

4.12 Quasi-static compression of a gas

Consider a thermally insulated ideal gas of particles confined within a container of volume V . The gas is initially at some absolute temperature T . Assume now that the volume of this container is very slowly reduced by moving a piston to a new position. Give qualitative answers to the following questions:

- What happens to the energy levels of each particle?
- Does the mean energy of a particle increase or decrease?
- Is the work done on the gas in reducing its volume positive or negative?
- Does the mean energy of a particle, measured above its ground state energy, increase or decrease?
- Does the absolute temperature of the gas increase or decrease?

(b) By simply averaging, derive an expression for the mean force F in terms of the mean energy $\bar{\epsilon}$ of a particle. Use the symmetry requirement that $\overline{K_x^2} = \overline{K_y^2} = \overline{K_z^2}$ when the gas is in equilibrium.

(c) Hence show that the mean pressure $\bar{p}$ exerted by the gas is given by

$$\bar{p} = \frac{2}{3} \bar{u} \quad (\text{ii})$$

where $\bar{u}$ is the mean energy per unit volume of the gas.

4.17 Pressure and energy density of electromagnetic radiation

Consider electromagnetic radiation (i.e., a gas of photons) enclosed in a box of edge lengths L_x , L_y , and L_z . Since a photon moves with the speed c of light, it is a *relativistic* particle. Hence its energy ϵ is related to its momentum $\hbar K$ by

$$\epsilon = \hbar K = \hbar(K_x^2 + K_y^2 + K_z^2)^{1/2} \quad (\text{i})$$

where the possible values of K_x , K_y , and K_z are again given by (3.13).

(a) Use this expression to calculate the force F_r exerted by a photon on the right wall of the container when the photon is in a given state r specified by n_x, n_y, n_z .

(b) By simply averaging, derive an expression for the mean force F in terms of the mean energy $\bar{\epsilon}$ of a photon. Use the symmetry argument that $\overline{K_x^2} = \overline{K_y^2} = \overline{K_z^2}$ when the radiation is in equilibrium with the container walls.

(c) Hence show that the mean pressure $\bar{p}$ exerted on the walls by the radiation is given by

$$\bar{p} = \frac{1}{3} \bar{u} \quad (\text{ii})$$

where $\bar{u}$ is the mean electromagnetic energy per unit volume of the radiation.

(d) Why is the constant of proportionality in (ii) equal to $\frac{1}{3}$ instead of being equal to the value $\frac{2}{3}$ derived in the preceding problem for a nonrelativistic gas?

4.18 Mean energy expressed in terms of partition function

Consider any system, no matter how complicated, in thermal equilibrium with a heat reservoir at the absolute temperature $T = (k\beta)^{-1}$. The probability that the system is in any one of its states r of energy E_r is then given by the canonical distribution (49). Obtain an expression for the mean energy E of this system. In particular, show that the arguments used in Sec. 4.7 are generally applicable and derive thus the very general result

$$E = - \frac{\partial \ln Z}{\partial \beta} \quad (\text{i})$$

Here

$$Z \equiv \sum_r e^{-\beta E_r} \quad (\text{ii})$$

represents a sum over all possible states of the system and is called the *partition function* of the system.

4.19 Mean pressure expressed in terms of partition function

Consider again the system described in Prob. 4.18. The system is in thermal equilibrium with a heat reservoir at the absolute temperature T , but may be arbitrarily complicated (e.g., it might be a gas, a liquid, or a solid). For the sake of simplicity, assume that the system is confined within a container in the shape of a rectangular parallelepiped with edge lengths L_x , L_y , and L_z . Show that the arguments used in Sec. 4.8 are generally applicable and establish thus the following very general results:

(a) Show that the mean force $\bar{F}$ exerted by the system on its right boundary wall can always be expressed in terms of the partition function Z of the system by the relation

$$\bar{F} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial L_x}. \quad (i)$$

Here Z is defined by the relation (ii) of the preceding problem.

(b) In the case of any isotropic system, the function Z does not depend on the individual dimensions L_x , L_y , and L_z , but is merely a function of the volume $V = L_x L_y L_z$ of the system. Show that (i) implies then that the mean pressure $\bar{p}$ exerted by the system can be expressed in the form

$$\bar{p} = \frac{1}{\beta} \frac{\partial \ln Z}{\partial V}. \quad (ii)$$

4.20 Partition function of an entire gas

Consider an ideal gas consisting of N monatomic molecules.

(a) Write down the expression for the partition function Z of this entire gas. By exploiting the properties of the exponential function, show that Z can be written in the form

$$Z = Z_0^N \quad (i)$$

where Z_0 is the partition function for a single molecule and was already calculated in Sec. 4.7.

(b) Use (i) to calculate the mean energy $\bar{E}$ of the gas by means of the general relation derived in Prob. 4.18. Show that the functional form of (i) implies immediately that $\bar{E}$ must be simply N times as large as the mean energy per molecule.

(c) Use (i) to calculate the mean pressure $\bar{p}$ of the gas by means of the general relation derived in Prob. 4.19. Show that the functional form of (i) implies again that $\bar{p}$ must be simply N times as large as the mean pressure exerted by a single molecule.

4.21 Mean energy of a magnetic moment

Consider a single spin $\frac{1}{2}$ in contact with a heat reservoir at the absolute temperature T . The spin has a magnetic moment μ_0 and is located in an external magnetic field B .

- (a) Calculate the partition function Z of this spin.
 (b) Using your result for Z , apply the general relation (i) of Prob. 4.18 to obtain the mean energy E of this spin as a function of T and B .
 (c) Verify that your expression for E satisfies the expression $E = -\bar{\mu}B$ where $\bar{\mu}$ is the value of the mean component of magnetic moment previously derived in Eq. (59).

4.22 Mean energy of a harmonic oscillator

A harmonic oscillator has a mass and spring constant which are such that its classical angular frequency of oscillation is equal to ω . In a quantum-mechanical description, such an oscillator is characterized by a set of discrete states having energies E_n given by

$$E_n = (n + \frac{1}{2})\hbar\omega. \quad (i)$$

The quantum number n which labels these states can here assume all the integral values

$$n = 0, 1, 2, 3, \dots \quad (ii)$$

A particular instance of a harmonic oscillator might, for example, be an atom vibrating about its equilibrium position in a solid.

Suppose that such a harmonic oscillator is in thermal equilibrium with some heat reservoir at the absolute temperature T . To find the mean energy E of this oscillator, proceed as follows:

(a) First calculate the partition function Z for this oscillator, using the definition (ii) of Prob. 4.18. (To evaluate the sum, note that it is merely a geometric series.)

(b) Apply the general relation (i) of Prob. 4.18 to calculate the mean energy of the oscillator.

(c) Make a qualitative sketch showing how the mean energy E depends on the absolute temperature T .

(d) Suppose that the temperature T is very small in the sense that $kT \ll \hbar\omega$. Without any calculation whatever, using only the energy levels of (i), what can you say about the value of E in this case? Does the result you obtained in (b) properly approach this limiting case?

(e) Suppose that the temperature T is very high so that $kT \gg \hbar\omega$. What then is the limiting value of the mean energy E obtained in (b)? How does it depend on T ? How does it depend on ω ?

*4.23 Mean rotational energy of a diatomic molecule

The kinetic energy of a diatomic molecule, rotating about an axis perpendicular to the line joining the two atoms, is classically given by

$$E = \frac{J^2}{2A} = \frac{J^2}{2A}$$

where J is the angular momentum and A is the moment of inertia of the mole-

cule. In a quantum-mechanical description, this energy can assume the discrete values

$$E_j = \frac{\hbar^2 j(j+1)}{2A} \quad (i)$$

where the quantum number j , which determines the *magnitude* of the angular momentum $\mathbf{J}$, can assume the possible values

$$j = 0, 1, 2, 3, \dots \quad (ii)$$

For each value of j , there are $(2j+1)$ distinct possible quantum states which correspond to the discrete possible spatial orientations of the angular momentum vector $\mathbf{J}$.

Suppose that the diatomic molecule is in a gas in equilibrium at the absolute temperature T . To calculate the mean energy of rotation of this diatomic molecule, proceed as follows:

(a) First calculate the partition function Z , using the definition (ii) of Prob. 4.18. (Be careful to remember that this is a sum containing a term for *each* individual state.) Assume that T is sufficiently large so that $kT \gg \hbar^2/2A$, a condition satisfied for most diatomic molecules at room temperature. Show that the sum Z can then be approximated by an integral, using $u = j(j+1)$ as a continuous variable.

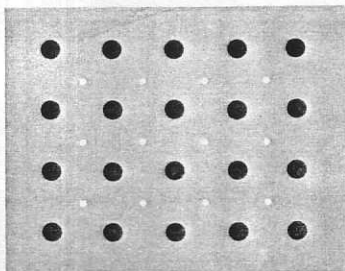
(b) Now apply the general relation (i) of Prob. 4.18 to calculate the mean rotational energy $\bar{E}$ of the diatomic molecule in this temperature range.

4.24 Number of interstitial atoms in a solid (approximate analysis)

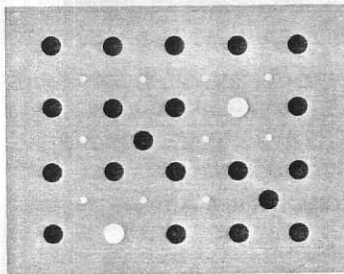
Consider a monatomic crystalline solid consisting of N atoms and maintained at an absolute temperature T . The atoms are ordinarily located at the normal lattice position indicated by the black circles in Fig. 4.17a. An atom may, however, also be located at one of the *interstitial positions* indicated by the white dots in this figure. If an atom is in such an interstitial position, its energy is larger by an amount ϵ than when it is in a normal position. When the absolute temperature is very low, all atoms will, therefore, be in normal positions. When the absolute temperature T is appreciable, however, this is no longer the case. Suppose that there are N atoms which can be located in N possible normal and N possible interstitial positions. The question of interest is then the following: At any absolute temperature T , what is the mean number $\bar{n}$ of atoms located in interstitial positions? An approximate way of solving the problem is the following:

(a) Focus attention first on an individual atom and consider that it can be in either one of only two particular positions—one normal, the other interstitial. This system can then be in only two possible configurations, call them A and B:

- (A): atom in normal position,
no atom in interstitial position.
- (B): no atom in normal position,
atom in interstitial position.



(a)



(b)

Fig. 4.17 In (a) all the atoms of the solid (indicated by black circles) are shown in their normal positions while the possible interstitial positions (indicated by white dots) are unoccupied. At higher temperatures some of the interstitial positions are also occupied, as shown in (b).

What is the ratio P_B/P_A of the probabilities P_B and P_A of encountering these two configurations?

(b) Focusing attention now on the whole solid, suppose that there are $\bar{n}$ atoms in interstitial positions. Then there must also be $\bar{n}$ atoms absent from normal positions. Since any one of the $\bar{n}$ empty normal positions can be combined with any one of the $\bar{n}$ occupied interstitial positions, a B configuration can arise in any of $\bar{n}^2$ possible ways. According to this argument, the probability P_B of encountering a single atom in the solid in a B configuration should be simply proportional to $\bar{n}^2$ if the empty normal and occupied interstitial positions are assumed to be distributed at random. Thus $P_B \propto \bar{n}^2$. Show by a similar argument that $P_A \propto (N - \bar{n})^2$.

(c) Combining the results of parts (a) and (b) and assuming the usual situation where $\bar{n} \ll N$, show that

$$\frac{\bar{n}}{N} = e^{-(1/2)\beta\epsilon}. \quad (i)$$

4.25 Number of interstitial atoms in a solid (exact analysis)

Consider the physical situation described in Prob. 4.24. To do the problem exactly, try to find the probability $P(n)$ that precisely n interstitial positions are occupied by atoms. There are then, of course, also n normal positions unoccupied by atoms.

(a) What is the probability of encountering a given situation where n interstitial atoms are distributed in *one* particular way and the n empty normal positions are also distributed in *one* particular way?

(b) In how many ways is it possible to distribute the n atoms over the N possible interstitial positions? In how many ways is it possible to distribute the n missing atoms over the N normal positions?

(c) Combining the results of parts (a) and (b), show that

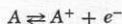
$$P(n) \propto \left[\frac{N!}{n!(N-n)!} \right]^2 e^{-\beta n \epsilon}. \quad (i)$$

(d) The probability $P(n)$ has a sharp maximum for some value $n = \bar{n}$. To find this value $\bar{n}$, consider $\ln P(n)$ and satisfy the condition $(\partial \ln P / \partial n) = 0$. Since all factorials are large, Stirling's approximation (M.10) is applicable. Show thus that, for $\bar{n} \ll N$, one obtains the result

$$\frac{\bar{n}}{N} = e^{-(1/2)\beta\epsilon}. \quad (ii)$$

*4.26 Thermal dissociation of an atom

An ideal gas of atoms is confined within a box with edge lengths L_x , L_y , and L_z . The whole system is in equilibrium at some absolute temperature T . The mass of an atom is M . An atom A can dissociate into a positive ion A^+ and an electron e^- according to the scheme



An ionization energy u is required to overcome the electron binding and achieve ionization.

Focusing attention on an individual atom, it can thus be in two possible configurations, call them U and D .

(U): The atom is undissociated. Its energy E is given by

$$E = \epsilon$$

where ϵ is the kinetic energy of its center of mass. The translational state of motion of the atom can, as usual, be specified by the set of quantum numbers $\{n_x, n_y, n_z\}$.

(D): The atom is dissociated into an electron of mass m and a positive ion of mass nearly equal to M (since $m \ll M$). The interaction between the ion and electron after dissociation can be assumed to be negligible. The total energy of the dissociated system consisting of two separate particles is then equal to

$$E = \epsilon^+ + \epsilon^- + u \quad (i)$$

where ϵ^+ is the kinetic energy of the ion, ϵ^- is the kinetic energy of the electron, and u is the ionization energy. The translational state of this dissociated system can then be specified by six quantum numbers, the set of quantum numbers $\{n_x^+, n_y^+, n_z^+\}$ of the ion and the set of quantum numbers $\{n_x^-, n_y^-, n_z^-\}$ of the electron.

(a) Using the canonical distribution find, to within a constant of proportionality C , the probability P_U that the atom is found among those of its states where it is in the undissociated configuration (U).

(b) Using the canonical distribution find, to within the same constant of proportionality C , the probability P_D that the atom is found among those of its states where it is in the dissociated configuration (D). [Note that the required sum over all the relevant states is the one already performed in calculating Z in Eq. (81) of Sec. 4.7.]

(c) Find the ratio P_D/P_U . How does it depend on the temperature T and the volume V ?

(d) Now consider the whole gas containing N atoms. Suppose that a mean number $\bar{n}$ of these are dissociated. Then there exist in the box $\bar{n}$ ions and $\bar{n}$ electrons, while there remain $(N - \bar{n})$ undissociated atoms. A dissociated configuration of an atom can then be realized in $\bar{n} \times \bar{n} = \bar{n}^2$ possible ways, an undissociated one in $(N - \bar{n})$ possible ways. By approximate reasoning similar to that of Prob. 4.24, one can then write

$$\frac{P_D}{P_U} = \frac{\bar{n}^2}{N - \bar{n}} \approx \frac{\bar{n}^2}{N}$$

if $\bar{n} \ll N$. Find thus an explicit expression for $(\bar{n}/N)$ in terms of the absolute temperature T and the density (N/V) of the gas.

(e) Ordinarily $kT \ll u$. Under these conditions would you expect most atoms to be dissociated or not?

(f) Suppose that $kT \ll u$, but that the volume of the box is made arbitrarily large while keeping the temperature T constant. Would most atoms then be dissociated or not? Give a simple physical explanation to account for this result.

(g) The inner part of the sun consists of very hot dense gases, while its outer corona is cooler and less dense. Studies of spectral lines of light from the sun indicate that an atom may be ionized in the corona, and yet be un-ionized in regions closer to the sun where the absolute temperature is much higher. How do you explain these observations?

*4.27 Thermal generation of a plasma

By heating a gas of atoms to sufficiently high temperature, one may generate a *plasma* consisting of an appreciable number of dissociated positive and negative charges. To study the practical possibility of this procedure, apply the results of Prob. 4.26 to cesium vapor. The cesium atom has a rather low ionization energy $u = 3.89$ electron volts and an atomic weight of 132.9.

(a) Express the degree of dissociation $\bar{n}/N$ of Prob. 4.26 in terms of T and the mean pressure $\bar{p}$ of the gas.

(b) Suppose that one heats cesium vapor to an absolute temperature 4 times as large as room temperature and maintains it at a pressure of 10^3 dynes/cm² (i.e., 10^{-3} of atmospheric pressure). Calculate the percentage of vapor ionized under these conditions.

4.28 Dependence of energy on temperature for an ideal gas

The number of states $\Omega(E)$ of an ideal gas of N monatomic atoms depends on the total energy E of the gas in a manner derived in Prob. 3.8. Use this result and the definition $\beta = \partial \ln \Omega / \partial E$ to derive a relation expressing the energy E as a function of the absolute temperature $T = (k\beta)^{-1}$. Compare your result with the expression for $E(T)$ derived in Sec. 4.7.

4.29 Dependence of energy on temperature for a spin system

The number of states $\Omega(E)$ of a system of N spins $\frac{1}{2}$, each having a magnetic moment μ_0 and located in a magnetic field B , has been calculated in Prob. 3.9.

(a) Use this result and the definition $\beta = (\partial \ln \Omega / \partial E)$ to derive a relation expressing the energy E of this system as a function of the absolute temperature $T = (k\beta)^{-1}$.

(b) Since the total magnetic moment M of this system is simply related to its total energy E , use the answer to part (a) to find an expression for M as a function of T and B . Compare this expression with the result derived for $\bar{M}_0$ in (61) and (59).

***4.30 Negative absolute temperature and heat flow in a spin system**

A system consists of N spins $\frac{1}{2}$, each having magnetic moment μ_0 and located in an external magnetic field B . The number of states $\Omega(E)$ of this system has already been calculated as a function of its total energy E in Prob. 3.9.

(a) Make an approximate sketch showing the behavior of $\ln \Omega$ as a function of E . Note that the lowest energy of the system is $E_0 = -N\mu_0 B$ and its highest energy is $+N\mu_0 B$, and that the curve is symmetric about the value $E = 0$.

(b) Use the curve of part (a) to make an approximate sketch showing β as a function of E . Note that $\beta = 0$ for $E = 0$.

(c) Use the curve of part (b) to make an approximate sketch showing the absolute temperature T as a function of E . What happens to T near $E = 0$? What is the sign of T for $E < 0$ and for $E > 0$?

(d) Since T suffers a discontinuity near $E = 0$, it is more convenient to work in terms of β . Show that $\partial\beta/\partial E$ is always negative. Hence prove that, when two systems are placed in thermal contact, heat is always absorbed by the system with the larger value of β . Note that this last statement is generally valid for all systems, irrespective of whether their absolute temperatures are positive or negative.

The very recent success with even consumer products involving fuzzy logic has caused a growing interest in the field of fuzzy sets and fuzzy control.

This book is written for scientists and engineers as well as logicians and mathematicians who are interested in fuzzy logic control and the foundations of fuzzy theory. The author, a specialist in the field for many

Microscopic Theory and Macroscopic Measurements

the use of notions and methods of many-valued logic. Thus, the book provides the essentials of a thorough discussion of the building processes with fuzzy sets, especially related to power systems and fuzzy control.

5.1	<i>Determination of the Absolute Temperature</i>	192
5.2	<i>High and Low Absolute Temperatures</i>	196
5.3	<i>Work, Internal Energy, and Heat</i>	200
5.4	<i>Heat Capacity</i>	206
5.5	<i>Entropy</i>	209
5.6	<i>Intensive and Extensive Parameters</i>	211
	<i>Summary of Definitions</i>	213
	<i>Important Relations</i>	213
	<i>Suggestions for Supplementary Reading</i>	213
	<i>Problems</i>	214

Chapter 6

Canonical Distribution in the Classical Approximation

6.1	<i>The Classical Approximation</i>	224
6.2	<i>Maxwell Velocity Distribution</i>	231
6.3	<i>Discussion of the Maxwell Distribution</i>	235
6.4	<i>Effusion and Molecular Beams</i>	240
6.5	<i>The Equipartition Theorem</i>	246
6.6	<i>Applications of the Equipartition Theorem</i>	248
6.7	<i>The Specific Heat of Solids</i>	250
	<i>Summary of Definitions</i>	256
	<i>Important Relations</i>	256
	<i>Suggestions for Supplementary Reading</i>	256
	<i>Problems</i>	257

Chapter 6 Canonical Distribution in the Classical Approximation

The canonical distribution (4.49) represents a simple result of fundamental importance and exceedingly great practical utility. As we showed in Chap. 4, it can be applied directly to calculate the equilibrium properties of the most diverse systems. Thus we illustrated specifically how it can be used to derive the magnetic properties of a system of spins or to calculate the pressure and specific heat of an ideal gas. We also examined several interesting applications in the problems at the end of Chap. 4. It would carry us too far afield to discuss the wide range of other important applications, a task that might easily fill several books. In the present chapter we do, however, want to show how some particularly simple and useful results follow immediately from the canonical distribution when the approximations of classical mechanics are applicable.

6.1 The Classical Approximation

We know that the quantum-mechanical description of a system of particles, under appropriate circumstances, can be approximated by a description in terms of classical mechanics. In this section we shall want to examine the following two questions: (i) Under what conditions can a statistical theory in terms of classical concepts be expected to be a valid approximation? (ii) If the approximation is permissible, how can the statistical theory be formulated in classical terms?

Validity of the classical approximation

The classical approximation can certainly *not* be valid if the absolute temperature is sufficiently low. Indeed, suppose that the typical thermal energy kT is less than (or comparable to) the average spacing ΔE between the energy levels of the system. It is then very significant that the possible energies of the system are quantized so as to be separated by discrete amounts. For example, the canonical distribution (4.49) implies that the probabilities of finding the system in a state of energy E or in a state of next higher possible energy $E + \Delta E$ are then very different. On the other hand, if $kT \gg \Delta E$, the probabilities vary very little from state to state. The fact that the possible energies

Quantization

$$\Delta E \sim kT$$

are discrete rather than continuous then becomes relatively unimportant and a classical description might become possible. The definite conclusion emerging from these comments is that

a classical description *cannot* be valid if

$$kT \leq \Delta E. \quad (1)$$

The classical approximation should certainly be valid, however, if quantum-mechanical effects can be shown to be of negligible importance. The fundamental quantum-mechanical limitation on the meaningful use of classical concepts is expressed by the Heisenberg uncertainty principle. This asserts that a simultaneous determination of a position coordinate q and its corresponding momentum p cannot be accomplished to infinite precision, but that these quantities are subject to minimum uncertainties of respective magnitudes Δq and Δp such that

$$\Delta q \Delta p \geq \hbar \quad (2)$$

where $\hbar \equiv h/2\pi$ is Planck's constant divided by 2π . Let us then examine the classical description of a system at a particular temperature. To be meaningful, this classical description must be able to consider a particle of the system localized within some typical minimum distance which we shall denote by s_0 . Furthermore, we shall denote by p_0 the typical momentum of the particle. If s_0 and p_0 are large enough so that

$$s_0 p_0 \gg \hbar,$$

the limitations imposed by the Heisenberg uncertainty principle should become of negligible importance and the classical description should, accordingly, be valid. Thus we are led to the conclusion that

a classical description should be valid

$$\text{if} \quad s_0 p_0 \gg \hbar, \quad (3a)$$

$$\text{i.e., if} \quad s_0 \gg \lambda_0. \quad (3b)$$

Here we have introduced the typical length λ_0 defined by

$$\lambda_0 \equiv \frac{\hbar}{p_0} = \frac{1}{2\pi} \frac{h}{p_0}. \quad (4)$$

This is just the de Broglie wavelength h/p_0 divided by 2π . The statement (3b) asserts thus merely that quantum effects ought to be negli-

Wave behavior

w.r. to being a Particle



(3b) $\rightarrow$ (Valid) $\leftarrow$ (1)

- Continuity of spaces and the (range of action of each space)
- Space (Particle) living inside space (spatial), only the effects of the particle space are revealed in the space in the neighbourhood, depending on its internal space itself.
- hence (3b), λ_0 is the result of transmitted effect of the internal space, to the external space.

ble if the minimum significant classical dimension s_0 is large compared to the de Broglie wavelength of the particle. The wave properties of the particle then clearly become unimportant.

The classical description

Suppose that a classical discussion of a system of particles is justified. Then the basic questions that must be examined are precisely the same as those which formed the starting point of our quantum-theoretical considerations at the beginning of Chap. 3. In particular, the first question that arises is the following: How does one specify the microscopic state of a system described in terms of classical mechanics?

Let us begin by considering the very simple case of a system consisting of a single particle moving in one dimension. The position of this particle can be described by a single coordinate, call it q . The complete specification of the system in classical mechanics then requires a knowledge of the coordinate q and of its corresponding momentum p .† (A simultaneous knowledge of q and p at any one time is classically possible. It is also required for a complete description so that values of q and p at any other time can be predicted uniquely in accordance with the laws of classical mechanics.) It is possible to represent the situation geometrically by drawing cartesian axes labeled by q and p as shown in Fig. 6.1. Specification of q and p is then equivalent to specifying a point in this two-dimensional space (commonly called phase space).

In order to describe the situation involving the continuous variables q and p so that the possible states of the particle are countable, it is convenient to follow the procedure of Sec. 2.6 by subdividing the ranges of the variables q and p into arbitrarily small discrete intervals. For example, one can choose fixed small intervals of size δq for the subdivision of q , and fixed small intervals of size δp for the subdivision of p . Phase space is thus subdivided into small cells of equal size and of two-dimensional "volume" (i.e., area)

$$\delta q \delta p = h_0$$

where h_0 is some small constant (having the dimensions of angular momentum). A complete description of the state of the particle can then be given by specifying that its coordinate lies in some particular

(Hilbert Space is a Complete Inner-product Space).

† If q denotes an ordinary cartesian coordinate and if no magnetic field is present, the momentum p is simply related to the velocity v of the particle of mass m by the proportionality $p = mv$. The description in terms of the momentum p rather than the velocity v is, however, valid in more general cases and is thus the one commonly used.

(description)
Continuity of Mechanics
(how) from QCD $\rightarrow$ MGR

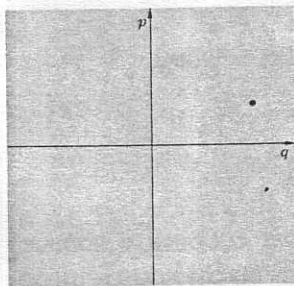


Fig. 6.1 Classical phase space for a single particle in one dimension.

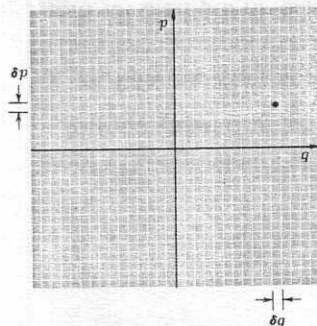


Fig. 6.2 The two-dimensional phase space of the preceding figure is here shown subdivided into equal cells of "volume" $\delta q \delta p = h_0$.

interval between q and $q + \delta q$ and that its momentum lies in some particular interval between p and $p + \delta p$, i.e., by specifying that the pair of numbers $\{q, p\}$ lies in some particular range. Geometrically this specifies that the point represented by $\{q, p\}$ lies in a particular cell of phase space.

Remark concerning the size of h_0

The specification of the state of the system is clearly more precise the smaller one chooses the size of the cells into which phase space has been divided, i.e., the smaller one chooses h_0 . This constant h_0 can be chosen arbitrarily small in a classical description. The correct quantum-mechanical description imposes, however, a limitation on the maximum accuracy that can be achieved in a simultaneous specification of a coordi-

nate q and its corresponding momentum p . Indeed, q and p can only be determined within uncertainties Δq and Δp whose order of magnitude satisfies the Heisenberg uncertainty principle $\Delta q \Delta p > \hbar$. A subdivision of phase space into cells of volume less than $\hbar$ is thus physically meaningless; i.e., a choice of $h_0 < \hbar$ would lead to a specification of the system more precise than is allowed by quantum theory.

The generalization of the previous discussion to an arbitrarily complex system is immediate. Such a system can be described by some set of f coordinates $q_1, \dots, q_f$ and f corresponding momenta $p_1, \dots, p_f$, i.e., by a total of $2f$ numbers. (As usual, the number f of independent coordinates needed for the description of the system is called the *number of degrees of freedom* of the system.) In order to deal with these continuous variables in a manner where the possible states of the system are countable, it is again convenient to subdivide the possible values of the i th coordinate q_i into fixed small intervals of magnitude δq_i , and the possible values of the i th momentum p_i into fixed small intervals of magnitude δp_i . For each i , the size of the subdivision interval can be chosen so that the product

$$\delta q_i \delta p_i = h_0 \quad (5)$$

where h_0 is some arbitrarily small constant of fixed magnitude independent of i . The state of the system can then be specified by stating that its coordinates and momenta are such that the set of values

$$\{q_1, q_2, \dots, q_f; p_1, p_2, \dots, p_f\} = (\underline{q}, \underline{p})$$

lies in a particular set of intervals. In a convenient geometrical interpretation this set of values can again be regarded as a "point" in a phase space of $2f$ dimensions where each cartesian axis is labeled

by one of the coordinates or momenta.[†] The subdivision into intervals thus divides this space into equal small cells of volume $(\delta q_1 \delta q_2 \cdots \delta q_f \delta p_1 \delta p_2 \cdots \delta p_f) = h_0^f$. The state of the system can then be described by specifying in which particular set of intervals (i.e., in which cell in phase space) the coordinates $q_1, q_2, \dots, q_f$ and momenta $p_1, p_2, \dots, p_f$ of the system actually lie. For simplicity each such set of intervals (or cell in phase space) can be labeled by some index r so that all these possible cells can be listed and enumerated in some convenient order $r = 1, 2, 3, \dots$. Our entire discussion can then be summarized by the observation that

the state of a system in classical mechanics can be described by specifying the particular cell r in phase space in which the coordinates and momenta of the system are found. (6)

The specification of the state of a system in classical mechanics is thus very similar to that in quantum mechanics, a cell in phase space in the classical description being analogous to a quantum state in the quantum-mechanical description. One distinction, however, is worth noting. In the classical case there exists an element of arbitrariness since the size of a cell in phase space (i.e., the magnitude of the constant h_0) can be chosen at will. On the other hand, in the quantum description a quantum state is an unambiguously defined entity (essentially because quantum theory involves Planck's constant $\hbar$ which has a unique value).

Classical statistical mechanics

The statistical description of a system in terms of classical mechanics now becomes completely analogous to that in quantum mechanics. The difference is one of interpretation: whereas the microstate of a system refers to a particular quantum state of a system in the quantum theory, it refers to a particular cell of phase space in the classical theory. When considering a statistical ensemble of systems, the basic postulates introduced in the classical theory are the same as the corresponding postulates (3.17) and (3.18) of quantum theory. In particular, the statement (3.19) of these postulates asserts in classical terms the fol-

[†] Except for the fact that it is not so readily visualized by our three-dimensional minds, this phase space is completely analogous to the two-dimensional phase space of Fig. 6.2.

lowing: If an isolated system is in equilibrium, it is found with equal probability in each one of its accessible states, i.e., in each one of its accessible cells in phase space.[†]

Example

In order to illustrate the classical notions in a very simple case, let us consider a single particle moving in one dimension under the influence of no forces, but confined within a box of length L . If we denote the position coordinate of this particle by x , the possible positions of the particle are then restricted by a condition of the form $0 < x < L$. The energy E of the particle of mass m is merely its kinetic energy so that

$$E = \frac{1}{2}mv^2 = \frac{1}{2}\frac{p^2}{m}$$

where v is the velocity and $p = mv$ is the momentum of the particle. Suppose that the particle is isolated and is thus known to have a constant energy in some small range between E and $E + dE$. Then its momentum must lie in some small range dp about the possible values $p = \pm\sqrt{2mE}$. The region of phase space accessible to this particle is then the one indicated by the dark areas

shown in Fig. 6.3. If phase space has been subdivided into small cells of equal size $\delta x \delta p = h_0$, this region contains a large number of such cells. These represent the accessible states in which the system can be found.

Suppose that the particle is known to be in equilibrium. Then the statistical postulate asserts that the particle is equally likely to be found with its coordinate x and momentum p in any one of the equal-size cells contained within the dark areas. This implies that the particle is as likely to have a momentum in the range dp near $+\sqrt{2mE}$ as in the range dp near $-\sqrt{2mE}$. It also implies that the position coordinate x of the particle is equally likely to lie anywhere within the length L of the box. For example, the probability that the particle is located in the left third of the box is $\frac{1}{3}$ since the number of accessible cells for which x lies in the range $0 < x < \frac{1}{3}L$ is one-third of the total number of accessible cells.

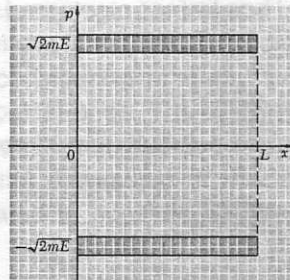


Fig. 6.3 Classical phase space for a single particle free to move in one dimension and confined within a box of length L . The particle, specified by a coordinate x and a momentum p , has an energy in the range between E and $E + \delta E$. The states accessible to the particle are indicated by the cells contained within the dark areas.

The preceding comments make it apparent that any general argument based upon the statistical postulates and the counting of states must remain equally valid in the classical description. In particular, it follows that the derivation of the canonical distribution in Sec. 4.5 remains applicable. If a system A , described classically, is in thermal equilibrium with a heat reservoir at the absolute temperature

[†] The statistical postulates can, with certain assumptions, be derived from the laws of classical mechanics in the same way that (3.17) and (3.18) could be based upon the laws of quantum mechanics. Such a classical derivation shows, incidentally, that a description in terms of coordinates and momenta (rather than coordinates and velocities) is the appropriate one in the most general case. In all the simple cases discussed in this book the distinction is trivial, however, since the momentum p and velocity v of a particle of mass m in these cases will always be related by the simple proportionality $p = mv$.

$T = (k\beta)^{-1}$, the probability P_r of finding this system in a particular state r of energy E_r is thus given by (4.49) so that

$$P_r \propto e^{-\beta E_r}. \quad (7)$$

Here the state r refers to a particular cell of phase space where the coordinates and momenta of A have particular values $\{q_1, \dots, q_f, p_1, \dots, p_f\}$. Correspondingly, the energy E_r of A denotes the energy E of this system when its coordinates and momenta have these particular values, i.e.,

$$E_r = E(q_1, \dots, q_f, p_1, \dots, p_f) \quad (8)$$

since the energy of A is a function of its coordinates and momenta.

It is convenient to express the canonical distribution (7) in terms of a probability *density* by proceeding in the general manner familiar from Sec. 2.6. Let us thus try to find the following probability

$$P(q_1, \dots, q_f, p_1, \dots, p_f) dq_1 \dots dq_f dp_1 \dots dp_f$$

(9)

≡ the probability that the system A in contact with the heat reservoir is found to have its first coordinate in the range between q_1 and $q_1 + dq_1, \dots$, its f th coordinate in the range between q_f and $q_f + dq_f$, its first momentum in the range between p_1 and $p_1 + dp_1, \dots$, and its f th momentum in the range between p_f and $p_f + dp_f$.

Here the ranges dq_i and dp_i are supposed to be small in the sense that the energy E of A does not vary appreciably when q_i changes by an amount dq_i or p_i changes by an amount dp_i . They are, however, supposed to be large compared to the intervals used in the subdivision of phase space, i.e., $dq_i \gg \delta q_i$ and $dp_i \gg \delta p_i$. The element of volume $(dq_1 \dots dq_f dp_1 \dots dp_f)$ of phase space thus contains many cells, each of volume $(\delta q_1 \dots \delta q_f \delta p_1 \dots \delta p_f) = h_0^f$. (See Fig. 6.4.) In each of these cells the energy of the system A , and hence also its probability (7), is then nearly the same. Hence the desired probability (9) is found simply by multiplying the probability (7) of finding A in a given cell of phase space by the total number $(dq_1 \dots dp_f)/h_0^f$ of such cells, i.e.,

$$P(q_1, \dots, p_f) dq_1 \dots dp_f \propto e^{-\beta E_r} \frac{dq_1 \dots dp_f}{h_0^f}$$

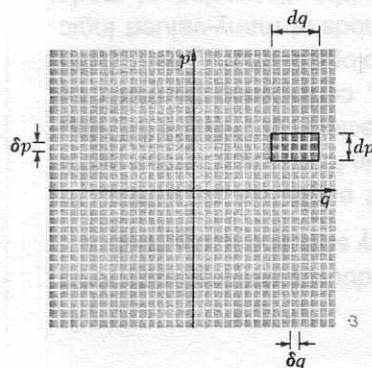


Fig. 6.4 A two-dimensional example of phase space subdivided into small cells of equal "volume" $\delta q \delta p = h_0$. The dark region indicates an element of volume having a size $dq dp$ and containing many cells.

$$\text{or } \boxed{\mathcal{P}(q_1, \dots, p_f) dq_1 \dots dp_f = C e^{-\beta E(q_1, \dots, p_f)} dq_1 \dots dp_f} \quad (10)$$

where C is merely some constant of proportionality (which includes the constant h_0^f). The value of this constant is, of course, determined by the normalization requirement that the sum of the probability (10) over all accessible coordinates and momenta of the system A be unity, i.e., that

$$\int \mathcal{P}(q_1 \dots p_f) dq_1 \dots dp_f = 1$$

where the integral extends over the entire region of phase space accessible to the system A . Thus it follows immediately that

$$C^{-1} = \int e^{-\beta E(q_1, \dots, p_f)} dq_1 \dots dp_f. \quad (11)$$

These general considerations will be illustrated in the next section by applying them to a simple case of great importance, that of a single molecule moving in three dimensions.

6.2 Maxwell Velocity Distribution

Consider an ideal gas confined within a container of volume V and in equilibrium at the absolute temperature T . This gas may consist of several different types of molecules. We shall assume that conditions are such that a classical treatment of the molecules in the gas is permissible. At the end of our discussion we shall examine the range of conditions within which such classical considerations can be expected to be valid. Let us think thus in classical terms and focus attention on any one of the molecules of the gas. This molecule then constitutes a distinct small system in thermal contact with a heat reservoir which consists of all the other molecules and has a temperature T . The canonical distribution is thus immediately applicable. Suppose, for the time being, that the molecule is monatomic. If one neglects any external force fields (such as gravity), the energy ϵ of this molecule then is simply its kinetic energy

$$\epsilon = \frac{1}{2} m v^2 = \frac{1}{2} \frac{\mathbf{p}^2}{m} \quad (12)$$

where $\mathbf{v}$ is the velocity and $\mathbf{p} = m\mathbf{v}$ is the momentum of the molecule of mass m . Here we have assumed that the gas is sufficiently dilute to be ideal; hence any potential energy of interaction with other mole-

cules is supposed to be negligible. The energy of the molecule anywhere within the container is thus independent of the position vector $\mathbf{r}$ of the molecule.

The state of the molecule is described classically in terms of the three position coordinates x, y, z of the molecule and its corresponding three momentum components p_x, p_y, p_z . We can then ask for the probability that the position of the molecule lies in the range between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ (i.e., that its x coordinate lies between x and $x + dx$, its y coordinate lies between y and $y + dy$, and its z coordinate lies between z and $z + dz$) and that simultaneously its momentum lies in the range between $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$ (i.e., that its x component of momentum lies between p_x and $p_x + dp_x$, its y component of momentum lies between p_y and $p_y + dp_y$, and its z component of momentum lies between p_z and $p_z + dp_z$). This range of position and momentum variables corresponds to a "volume" of phase space of size $(dx dy dz dp_x dp_y dp_z) \equiv d^3\mathbf{r} d^3\mathbf{p}$. Here we have introduced the convenient abbreviations

$$d^3\mathbf{r} \equiv dx dy dz$$

$$\text{and} \quad d^3\mathbf{p} \equiv dp_x dp_y dp_z \quad (13)$$

for an element of volume of real space and an element of volume of momentum space, respectively. Applying the canonical distribution (10), we then immediately obtain for the desired probability $\mathcal{P}(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ that the molecule has a position between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ and a momentum between $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$ the result

$$\mathcal{P}(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} \propto e^{-\beta(p^2/2m)} d^3\mathbf{r} d^3\mathbf{p} \quad (14)$$

where $\beta = (kT)^{-1}$. Here we have used the expression (12) for the energy of the molecule and have written $p^2 = p^2$. Equivalently we can express this result in terms of the velocity $\mathbf{v} = \mathbf{p}/m$ of the molecule to find the probability $\mathcal{P}'(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v}$ that the molecule has a position between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ and a velocity between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$. Thus

$$\mathcal{P}'(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} \propto e^{-(1/2)\beta m v^2} d^3\mathbf{r} d^3\mathbf{v} \quad (15)$$

where $d^3\mathbf{v} \equiv dv_x dv_y dv_z$ and where $v^2 = \mathbf{v} \cdot \mathbf{v}$.

The probability (15) is a very general result which provides detailed information about the position and velocity of any molecule in the gas. It allows us to deduce readily a variety of more special results. For

1 example, we may ask how many molecules have a velocity in a specified
 2 range. Or more generally, if the gas consists of a mixture of different
 kinds of molecules having different masses (e.g., helium and argon
 molecules), we may ask how many molecules of a given kind have a
 3 velocity in any specified range. Focusing attention on the molecules
 of a particular kind, we can thus try to find

$$f(\mathbf{v}) d^3\mathbf{v} \equiv \text{mean number of molecules (of the specified kind), per unit volume, which have a velocity between } \mathbf{v} \text{ and } \mathbf{v} + d\mathbf{v}. \quad (16)$$

i.e. we have two conditions

1 Since the N molecules of the ideal gas move independently without
 appreciable mutual interaction, the gas constitutes a statistical en-
 semble of molecules of which a fraction given by the probability (15)
 have a position between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ and a velocity between $\mathbf{v}$ and
 2 $\mathbf{v} + d\mathbf{v}$. The mean number $f(\mathbf{v}) d^3\mathbf{v}$ of (16) is thus simply obtained by
 multiplying the probability (15) by N , the total number of molecules
 of this kind, and dividing by the volume element $d^3\mathbf{r}$. Thus

$$f(\mathbf{v}) d^3\mathbf{v} = \frac{N \mathcal{P}'(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v}}{d^3\mathbf{r}}$$

or

$$f(\mathbf{v}) d^3\mathbf{v} = C e^{-(1/2)\beta m v^2} d^3\mathbf{v} \quad (17)$$

where C is a constant of proportionality and $\beta \equiv (kT)^{-1}$. The result
 3 (17) is called the *Maxwell velocity distribution* since it was first derived
 by Maxwell in 1859 (by the use of less general arguments).

4 1 Note that the probability $\mathcal{P}'$ of (15) [or the mean number f of (17)]
 does not depend on the position $\mathbf{r}$ of the molecule. This result must,
 5 of course, be true by virtue of symmetry considerations since a mole-
 cule can have no preferred position in space in the absence of external
 6 force fields. Note also that $\mathcal{P}'$ (or f) depends only on the magnitude
 of $\mathbf{v}$ and not on its direction; i.e.,

$$f(\mathbf{v}) = f(v) \quad (18)$$

1 where $v = |\mathbf{v}|$. Again this is obvious by symmetry, since there is no
 preferred direction in a situation where the container (and thus also
 the center of mass of the whole gas) is considered to be at rest.

Determination of the constant C

The constant C can be determined by the requirement that the sum of (17) over all possible velocities must yield the total mean number n of molecules (of the kind considered) per unit volume. Thus

$$C \int e^{-(1/2)\beta m v^2} d^3v = n \quad (19)$$

$$\text{or } C \iiint e^{-(1/2)\beta m (v_x^2 + v_y^2 + v_z^2)} dv_x dv_y dv_z = n.$$

$$\text{Thus } C \iiint e^{-(1/2)\beta m v_x^2} e^{-(1/2)\beta m v_y^2} e^{-(1/2)\beta m v_z^2} dv_x dv_y dv_z = n,$$

$$\text{or } C \int_{-\infty}^{\infty} e^{-(1/2)\beta m v_x^2} dv_x \int_{-\infty}^{\infty} e^{-(1/2)\beta m v_y^2} dv_y \int_{-\infty}^{\infty} e^{-(1/2)\beta m v_z^2} dv_z = n.$$

Each of these integrals has, by (M.23), the same value

$$\int_{-\infty}^{\infty} e^{-(1/2)\beta m v_x^2} dv_x = \left(\frac{\pi}{\frac{1}{2}\beta m}\right)^{1/2} = \left(\frac{2\pi}{\beta m}\right)^{1/2}.$$

$$\text{Hence } C = n \left(\frac{\beta m}{2\pi}\right)^{3/2} \quad (20)$$

$$\text{and } \boxed{f(\mathbf{v}) d^3v = n \left(\frac{\beta m}{2\pi}\right)^{3/2} e^{-(1/2)\beta m v^2} d^3v.} \quad (21)$$

Validity of the results for polyatomic molecules

Suppose that the gas under consideration contains molecules which are *not* monatomic. Under the conditions contemplated in the preceding paragraphs, the motion of the center of mass of such a molecule can still be treated by the classical approximation, although the intramolecular motion of rotation and vibration about its center of mass must ordinarily be discussed in terms of quantum mechanics. The state of the molecule can then be described by the position $\mathbf{r}$ and momentum $\mathbf{p}$ of its center of mass, and by specifying the particular quantum state s describing its intramolecular motion. The energy of the molecule is then

$$\epsilon = \frac{\mathbf{p}^2}{2m} + \epsilon_s^{(i)} \quad (22)$$

where the first term on the right is the kinetic energy of its center-of-mass motion and the second term is its intramolecular energy of rotation and vibration in the state s . The canonical distribution allows us to write down immediately an expression for the probability $P_s(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ that the molecule is found

in the state where its center-of-mass position is between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$, its center-of-mass momentum is between $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$, and its intramolecular motion is specified by s . Thus

$$\begin{aligned} P_s(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p} &\propto e^{-\beta[p^2/2m + \epsilon_s^{(i)}]} d^3\mathbf{r} d^3\mathbf{p} \\ &\propto e^{-\beta p^2/2m} d^3\mathbf{r} d^3\mathbf{p} e^{-\beta \epsilon_s^{(i)}}. \end{aligned} \quad (23)$$

To find the probability $P(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ that the center of mass has a position between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ and a momentum between $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$, *irrespective of the molecule's state of intramolecular motion*, it is only necessary to sum (23) over all possible intramolecular states s . But since the expression (23) is simply a product of two factors, the sum over all possible values of the second factor merely yields some constant multiplying the first factor. The result (23) reduces thus to an expression of the form (14) describing the center of mass of the molecule. Hence (15) and the Maxwell velocity distribution (17) are very general results which remain valid also in describing the center-of-mass motion of a polyatomic molecule in a gas.

6.3 Discussion of the Maxwell Distribution

- 1 The Maxwell velocity distribution (17) allows us to deduce immediately some related velocity distributions, in particular the distribution of
- 2 molecular speeds in a gas. As we shall see later, some of these results can be tested very directly by experiment. Let us explore some of these consequences of the Maxwell distribution and then examine the conditions under which the Maxwell distribution may be expected to be valid.

Distribution of a velocity component

- 1 Suppose that we focus attention on the component of a molecule's velocity along a particular direction, say the x direction. Then we are concerned with the following quantity describing a given kind of molecule,

$g(v_x) dv_x \equiv$ the mean number of molecules, per unit volume, which have an x component of velocity in the range between v_x and $v_x + dv_x$ (irrespective of the values of the other velocity components).

- 2 We obtain this number by simply adding all the molecules which have an x component of velocity in this range. Thus

$$g(v_x) dv_x = \int_{(v_y)} \int_{(v_z)} f(\mathbf{v}) d^3\mathbf{v}$$

where the sums (i.e., integrations) extend over all the possible y and z velocity components of the molecules. Hence (17) yields

$$\begin{aligned} g(v_x) dv_x &= C \int_{(v_y)} \int_{(v_z)} e^{-(1/2)\beta m(v_x^2 + v_y^2 + v_z^2)} dv_y dv_z \\ &= C e^{-(1/2)\beta m v_x^2} dv_x \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(1/2)\beta m(v_y^2 + v_z^2)} dv_y dv_z, \end{aligned}$$

or

$$g(v_x) dv_x = C' e^{-(1/2)\beta m v_x^2} dv_x \quad (24)$$

since the integration over all values of v_y and v_z gives merely some constant which can be absorbed into C' , a new constant of proportionality.[†] The constant C' can again be determined by the requirement that the total mean number of molecules per unit volume is properly equal to n , i.e., by the condition

$$\int_{-\infty}^{\infty} g(v_x) dv_x = C' \int_{-\infty}^{\infty} e^{-(1/2)\beta m v_x^2} dv_x = n.$$

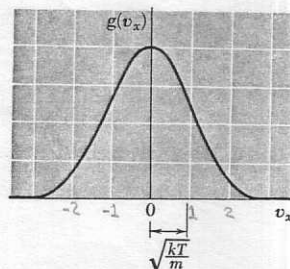


Fig. 6.5 Maxwell distribution showing the mean number $g(v_x) dv_x$ of molecules, per unit volume, having an x component of velocity between v_x and $v_x + dv_x$.

[†] Note that (24) is a simple Gaussian distribution of the kind discussed in Appendix A.1.

This gives
$$C' = n \left(\frac{\beta m}{2\pi} \right)^{1/2}. \quad (25)$$

The result (24) shows that the velocity component v_x is distributed symmetrically about the value $v_x = 0$. The *mean* value of any velocity component of a molecule must therefore always vanish, i.e.,

$$\bar{v}_x = 0. \quad (26)$$

- 1 This is physically clear by symmetry since the x component of velocity of a molecule is as likely to be positive as negative. Mathematically
2 this result follows from the definition of the average†

$$\bar{v}_x \equiv \frac{1}{n} \int_{-\infty}^{\infty} g(v_x) v_x dv_x.$$

Here the integrand is an odd function of v_x (i.e., reverses its sign when v_x reverses sign) because $g(v_x)$ is an even function of v_x (i.e., remains unchanged under this operation since it depends only on v_x^2). Thus contributions to the integrand from $+v_x$ and $-v_x$ cancel each other.

Note that $g(v_x)$ has its maximum value when $v_x = 0$ and decreases rapidly as $|v_x|$ increases. It becomes negligibly small when $|\beta m v_x^2| \gg 1$; that is,

$$\text{if } |v_x| \gg (kT/m)^{1/2}, \quad g(v_x) \rightarrow 0. \quad (27)$$

The distribution $g(v_x)$ thus becomes increasingly sharply peaked near $v_x = 0$ if the absolute temperature T is reduced. This merely reflects the fact that the mean kinetic energy of a molecule becomes increasingly small as $T \rightarrow 0$.

Needless to say, exactly similar results hold for the velocity components v_y and v_z since all velocity components are, by the symmetry of the situation, completely equivalent.

Distribution of molecular speeds

Considering a given kind of molecule, let us now investigate the quantity

$F(v) dv \equiv$ the mean number of molecules, per unit volume, which have a speed $v \equiv |v|$ in the range between v and $v + dv$.

- 1 We can obtain this number by adding all molecules having speeds in this range, irrespective of the *directions* of their velocity. Thus

$$F(v) dv = \int f(v) d^3v \quad (28)$$

† The average can here be written as an integral in the manner of Eq. (2.78).

where the prime on the integral indicates that the integration is over all velocities satisfying the condition that

$$v < |\mathbf{v}| < v + dv,$$

i.e., over all velocity vectors which terminate in velocity space within a spherical shell of inner radius v and outer radius $v + dv$. Since dv is infinitesimal and $f(\mathbf{v})$ depends only on the magnitude of $\mathbf{v}$, the function $f(\mathbf{v})$ has essentially the constant value $f(v)$ over the entire domain of integration of (28) and can thus be taken outside the integral. The remaining integral represents merely the volume in velocity space of a spherical shell of radius v and thickness dv , a volume equal to the area $4\pi v^2$ of this shell multiplied by its thickness dv . Hence (28) becomes simply

$$F(v) dv = 4\pi f(v) v^2 dv. \quad (29)$$

Using (17) this becomes explicitly

$$F(v) dv = 4\pi C e^{-(1/2)\beta m v^2} v^2 dv \quad (30)$$

where C is given by (20). The relation (30) is the Maxwell distribution of speeds. Note that it has a maximum for the same reason that is responsible for the maxima encountered in our general discussion of statistical mechanics. As v increases, the exponential factor *decreases*, but the volume of phase space available to the molecule is proportional to v^2 and *increases*; the net result is a gentle maximum.

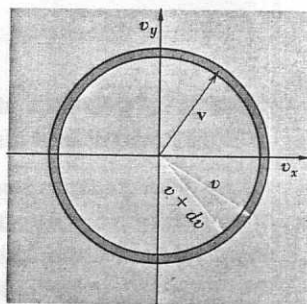


Fig. 6.6 Velocity space indicated in two dimensions, the v_z axis pointing out of the paper. The spherical shell contains all molecules having a velocity $\mathbf{v}$ such that $v < |\mathbf{v}| < v + dv$.

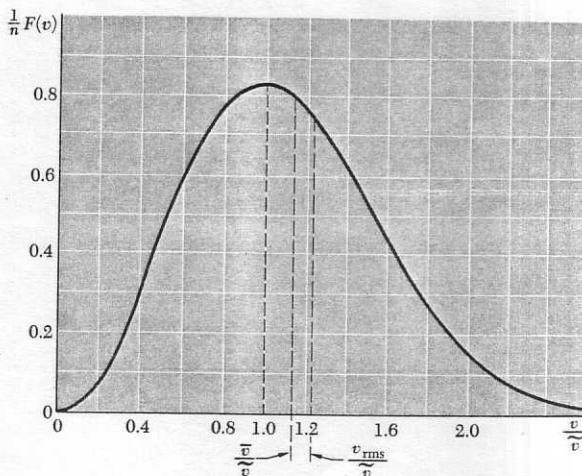


Fig. 6.7 Maxwell distribution showing the mean number $F(v) dv$ of molecules, per unit volume, having a speed between v and $v + dv$. The speed v is here expressed in terms of the most probable speed $\bar{v} = (2kT/m)^{1/2}$. Also shown are the mean speed $\bar{v}$ and the root-mean-square speed $v_{rms} \equiv (v^2)^{1/2}$.

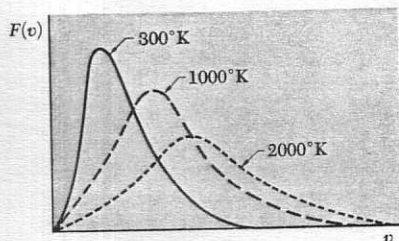


Fig. 6.8 Behavior of the Maxwell distribution of molecular speeds as a function of temperature.

Of course, if $F(v) dv$ is summed over all possible speeds $v = |\mathbf{v}|$, the result must again be the total mean number n of molecules per unit volume, i.e.,

$$\int_0^\infty F(v) dv = n. \quad (31)$$

The lower limit of the integral reflects the fact that, by its definition, the speed $v \equiv |\mathbf{v}|$ of a molecule cannot be negative.

A plot of $F(v)$ as a function of the speed v is shown in Fig. 6.7. The particular speed $v = \bar{v}$ where $F(v)$ has its maximum value is called the *most probable speed*. It can be found by setting $dF/dv = 0$. By using (30), this condition becomes

$$(-\beta m v e^{-(1/2)\beta m v^2})v^2 + e^{-(1/2)\beta m v^2}(2v) = 0$$

so that

$$\bar{v} = \sqrt{\frac{2}{\beta m}} = \sqrt{\frac{2kT}{m}}. \quad (32)$$

Let us, for example, consider nitrogen (N_2) gas at room temperature so that $T \approx 300^\circ\text{K}$. Since the molecular weight of N_2 is 28 and Avogadro's number is 6×10^{23} molecules/mole, the mass of a N_2 molecule is $m \approx 28/(6 \times 10^{23}) \approx 4.6 \times 10^{-23}$ gm. Hence (32) yields for the most probable speed of such a N_2 molecule

$$\bar{v} \approx 4.2 \times 10^4 \text{ cm/sec} = 420 \text{ m/sec}, \quad (33)$$

a number of the order of the speed of sound in the gas.

Validity of the classical discussion of a gas

Now let us examine under what conditions our classical discussion of the ideal gas, and hence also the Maxwell velocity distribution, can be expected to be valid. Our criterion of validity is the condition (3) following from the Heisenberg uncertainty principle. If the condition (3) is satisfied, the classical description ought to be adequate since it makes no statements that might violate limitations imposed by quantum ideas.

Since we are merely interested in typical orders of magnitude, we can be content with approximate estimates of the pertinent quantities in (3). The typical magnitude p_0 of the momentum of a molecule of mass m in a gas at temperature T can be found from the most probable speed $\bar{v}$ of such a molecule. Thus, by (32),

$$p_0 \approx m\bar{v} = \sqrt{2mkT}.$$

The corresponding typical de Broglie wavelength λ_0 of the molecule

is then

$$\lambda_0 \equiv \frac{\hbar}{p_0} \approx \frac{\hbar}{\sqrt{2mkT}}. \quad (34)$$

The classical description considers the molecules as distinguishable particles traveling along well-defined trajectories. This point of view ought certainly to be valid if there exist no quantum mechanical limitations preventing a molecule from being localized within a distance no greater than the typical separation s_0 between nearest molecules. This requires, in accordance with (3), that

$$s_0 \gg \lambda_0. \quad (35)$$

(The quantum-mechanical discussion shows that quantum effects *do* indeed become significant when the condition (35) is violated, precisely because the essential indistinguishability of the molecules is then of paramount importance). To estimate the typical separation s_0 between nearest molecules, imagine each molecule to be at the center of a little cube of side s_0 , these cubes filling the volume V available to the gas consisting of N molecules. Then

$$s_0^3 N = V$$

or

$$s_0 = \left(\frac{V}{N}\right)^{1/3} = n^{-1/3} \quad (36)$$

where $n \equiv N/V$ is the number of molecules per unit volume. The condition (35) for the validity of the classical approximation thus becomes

$$\frac{\lambda_0}{s_0} \approx \hbar \frac{n^{1/3}}{\sqrt{2mkT}} \ll 1. \quad (37)$$

This shows that the classical approximation ought to be applicable if the gas is sufficiently dilute so that n is small, if the temperature T is sufficiently high, and if the mass m of a molecule is not too small.

Numerical estimates

To estimate the typical magnitudes numerically, consider helium (He) gas at room temperature and atmospheric pressure (760 mm Hg). The relevant parameters are then:

mean pressure $\bar{p}$
 $= 760 \text{ mm Hg} \approx 10^6 \text{ dynes/cm}^2$;

temperature T
 $\approx 300^\circ\text{K}$; hence $kT \approx 4.1 \times 10^{-14} \text{ erg}$;

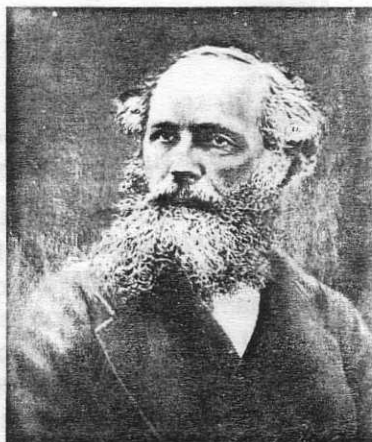
molecular mass m

$$= \frac{4}{6 \times 10^{23}} \approx 6.6 \times 10^{-24} \text{ gm.}$$

The equation of state of an ideal gas gives

$$n = \frac{\bar{p}}{kT} = 2.5 \times 10^{19} \text{ molecules/cm}^3.$$

Hence (34) and (36) yield the estimates



J. Clerk Maxwell

Fig. 6.9 James Clerk Maxwell (1831–1879). Although best known for his fundamental work in electromagnetic theory, Maxwell also contributed significantly to macroscopic thermodynamics and to the atomic theory of gases. He derived the molecular velocity distribution in 1859. After spending the earlier part of his career at the University of Aberdeen in Scotland, Maxwell became professor at Cambridge University in 1871. (From G. Holton and D. Roller, "Foundations of Modern Physical Science," Addison-Wesley Publishing Co., Inc., Cambridge, Mass., 1958. By permission of the publishers.)

$$\lambda_0 \approx 0.14 \text{ \AA}$$

and

$$s_0 \approx 33 \text{ \AA}$$

where $1 \text{ \AA} = 10^{-8} \text{ cm}$. Here the condition (35) is well satisfied and the classical approximation ought to be very good. Most gases have larger molecular weights and thus smaller de Broglie wavelengths; the condition (35) is then even better satisfied.

On the other hand, consider the conduction electrons in a typical metal such as copper. In first approximation, interactions between these electrons can be neglected so that they can be treated as an ideal gas. But the numerical values of the significant parameters are then quite different. The mass of an electron is very small, about 10^{-27} gm or 7300 times less than that of a He atom. This makes the de Broglie wavelength of the

electron much longer,

$$\lambda_0 \approx (0.14) \times \sqrt{7300} \approx 12 \text{ \AA}.$$

Furthermore, since there is about one conduction electron per atom in the metal and since a typical interatomic spacing is about $2 \text{ \AA}$,

$$s_0 \approx 2 \text{ \AA}.$$

The distance between particles is thus much smaller than in the case of He gas, i.e., the electrons in a metal form a very dense gas. These estimates show that the condition (35) is not satisfied by the electrons in a metal. Hence there exists no justification for discussing such electrons by classical statistical mechanics. In fact, it is then essential to give a completely quantum mechanical treatment which takes into account the Pauli exclusion principle obeyed by the electrons.

6.4 *Effusion and Molecular Beams*

Consider a gas in equilibrium inside some container. Suppose that a small hole of diameter D (or a narrow slit of width D) is now made in one of the walls of this container. If the hole is sufficiently small, the equilibrium of the gas inside the container should be disturbed to a negligible extent. The few molecules escaping through the hole into a vacuum surrounding the container should then constitute a representative sample of the molecules of the gas in its equilibrium state. Indeed, the molecules that have thus escaped can be collimated by slits to form a well-defined beam and, since they are few in number, their mutual interaction in this beam is insignificant. The molecules in such a beam can then be studied very effectively with two possible aims: (i) It may be of interest to study the properties of the molecules in the gas in equilibrium inside the container. For example, one may want to check whether the molecules in the container have velocities distributed in accordance with the predictions of the Maxwell distribution. (ii) It may be of interest to study the properties of essentially isolated molecules or atoms in order to investigate fundamental atomic or nuclear properties. Several Nobel prizes attest to the fruitfulness of this technique. We need only mention the fundamental experiments of Stern and Gerlach which led to the discovery of the spin and

associated magnetic moment of the electron, those of Rabi and co-workers which led to precision measurements of nuclear magnetic moments, and those of Kusch and Lamb which helped to lead to our modern understanding of the quantum theory of electromagnetic interactions.†

How small must be the size D of the hole so that the equilibrium of the gas inside the container is disturbed to a negligible extent? The hole must be sufficiently small so that the few molecules which are in the vicinity of the hole (and thus manage to escape through it) do not affect appreciably the large number of molecules in the remainder of the gas. This condition is satisfied if, during the time that a molecule spends in the vicinity of the hole, it suffers practically no collisions with other molecules. But if the mean speed of a molecule is $\bar{v}$, the time that a molecule spends near the hole is of the order of $D/\bar{v}$. On the other hand, the time between successive collisions of a molecule with other molecules is of the order of $l/\bar{v}$ if l denotes the mean free path of a molecule in the gas.‡ The preceding condition is thus equivalent to the statement that

$$\frac{D}{\bar{v}} \ll \frac{l}{\bar{v}}$$

or

$$D \ll l. \quad (38)$$

If this condition is satisfied, the molecules inside the container remain essentially in equilibrium (although their total number decreases slowly) and the escape of the molecules through the hole is said to constitute *effusion*.

Remark

The situation is appreciably different if $D \gg l$ so that molecules suffer frequent collisions with each other near the hole. When some molecules emerge through the hole (as in Fig. 6.10), the molecules behind them are then appreciably affected. They no longer continue colliding with the molecules on the right which have just escaped through the hole, but they still suffer frequent collisions with

the molecules on the left. These collisions cause the molecules near the hole to experience an unbalanced force to the right and thus to acquire a net velocity in the direction toward the hole. The resultant motion of all these molecules moving together as a group is then analogous to the flow of water through the hole of a tank. In this case one has not effusion, but *hydrodynamic flow*.

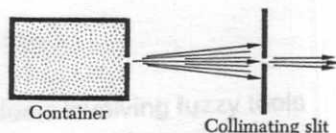


Fig. 6.10 Formation of a molecular beam by molecules escaping through a small hole in a container.

† A good and readable review of molecular beam experiments can be found in an article by O. R. Frisch in *Sci. American*, vol. 212, p. 58 (May 1965).

‡ As discussed in Sec. 1.6, the mean free path is defined as the mean distance traveled by a molecule in the gas before it collides with some other molecule.

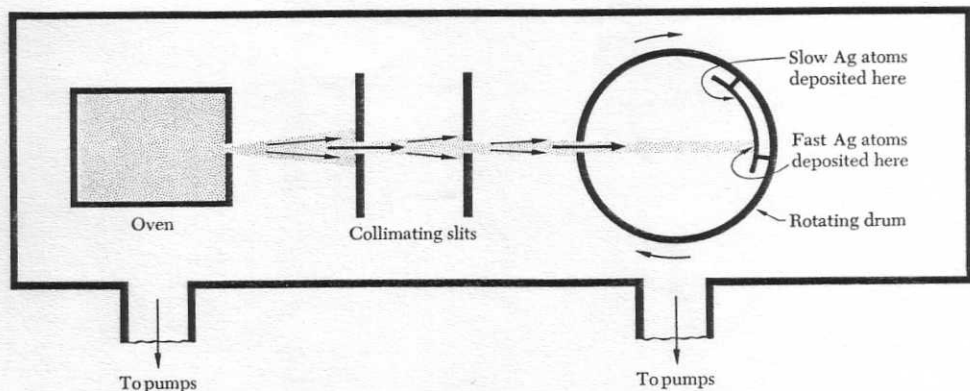


Fig. 6.11 A molecular-beam apparatus for studying the velocity distribution of silver (Ag) atoms. The Ag atoms stick to the drum surface upon impact.

If the hole is small enough so that the condition (38) is satisfied, the equilibrium of the gas is not affected appreciably by the presence of the hole. Accordingly the mean number $\mathcal{J}_0$ of molecules escaping per unit time through the hole is the same as the mean total number of molecules that would, per unit time, strike the area occupied by the hole if this hole had never been made. Hence $\mathcal{J}_0$ is simply given by the approximate expression (1.18) obtained in Sec. 1.6, i.e.,

$$\mathcal{J}_0 \approx \frac{1}{4} n \bar{v} \quad (39)$$

where n is the mean number of molecules per unit volume and $\bar{v}$ is their mean speed.[†] If one wants to focus attention only on those molecules having a speed between v and $v + dv$, the mean number $\mathcal{J}(v) dv$ of such molecules escaping through the hole per unit time is, analogously to (39), approximately given by

$$\mathcal{J}(v) dv \approx \frac{1}{4} [F(v) dv] v \quad (40)$$

where $F(v) dv$ denotes the mean number of molecules having a speed between v and $v + dv$. Using the Maxwell distribution of speeds (30), one then obtains the proportionality

$$\mathcal{J}(v) dv \propto v^3 e^{-(1/2)\beta m v^2}. \quad (41)$$

The last factor v in (40) expresses merely the fact that a fast molecule is more likely to escape through the hole than a slow one.

By measuring the relative number of molecules having various speeds in the molecular beam emerging from the hole in the container,

[†] An exact calculation yields, instead of (39), the result $\mathcal{J}_0 = \frac{1}{4} n \bar{v}$. See Appendix A.4 where the exact calculation is discussed.

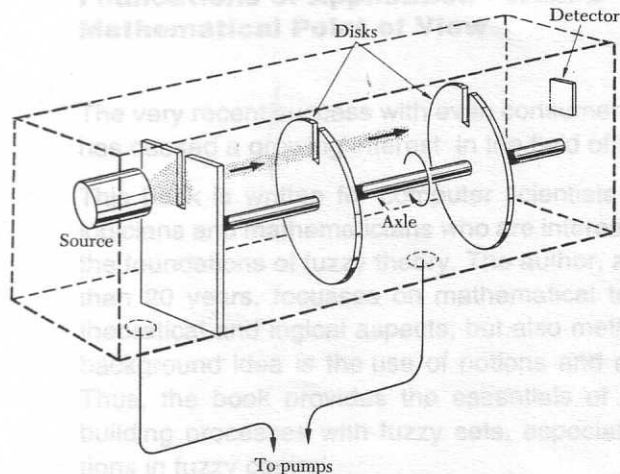


Fig. 6.12 A molecular-beam apparatus for studying the molecular velocity distribution by a velocity selector. By the time the beam of molecules reaches the second disk, the slot in this disk has ordinarily moved to a position such that the beam cannot pass through this disk. An exception occurs only if this disk has rotated by one revolution (or an integral number of revolutions) in the time required by the molecules to traverse the distance between the two disks. (A more effective velocity selector results if more than two similar disks are mounted on the same axle.)

one can test the prediction (41) and thus the Maxwell distribution upon which it is based. One experimental arrangement for achieving this purpose is shown in Fig. 6.11. Here silver is heated in an oven to generate a gas of silver (Ag) atoms, some of which emerge through a narrow slit to form an atomic beam. A hollow cylindrical drum, which has a slit in it and which rotates rapidly about its axis, is located in front of the beam. When the Ag atoms enter the slit in the drum, they require different times to reach the opposite side of the drum, a fast atom requiring less time than a slow one. Thus, since the drum is rotating, Ag atoms with different speeds strike the inside surface of the drum at different places and then stick to it. A subsequent measurement of the thickness of the deposited silver layer as a function of distance along the inside drum surface thus provides a measurement of the atomic velocity distribution.

A more accurate method for determining the velocity distribution uses a device which selects molecules having a particular velocity. (See Fig. 6.12. The method is analogous to the toothed-wheel method used by Fizeau to measure the speed of light.) In this method the molecular beam emerges from a hole and is detected at the other end of the apparatus. The velocity selector, which is placed between the source and the detector, consists in the simplest case of a pair of disks mounted on a common axle that can be rotated with known angular velocity. Both disks are identical and each has a slot cut into its periphery. The rotating disks act thus as two shutters which are alternately opened

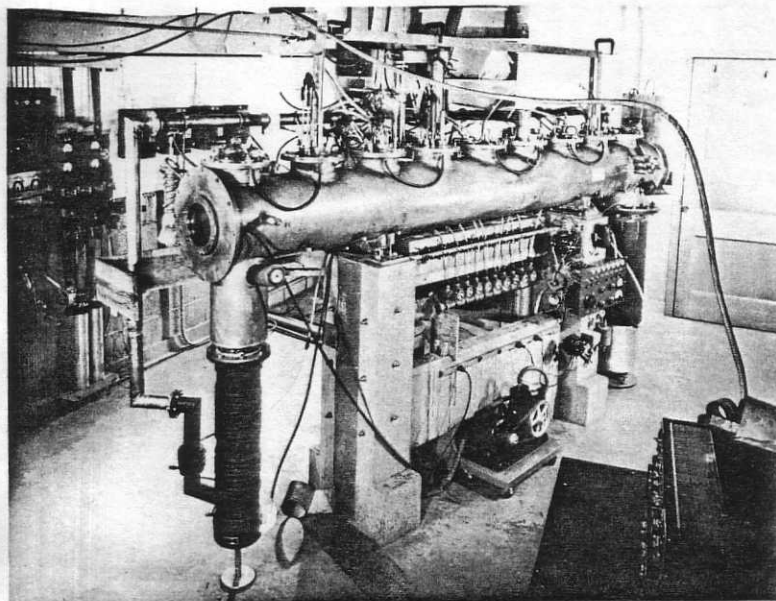


Fig. 6.13 Photograph of a modern molecular-beam apparatus used to study the properties of hydrogen molecules and atoms. (Photo-

graph by courtesy of Professor Norman F. Ramsey, Harvard University.)

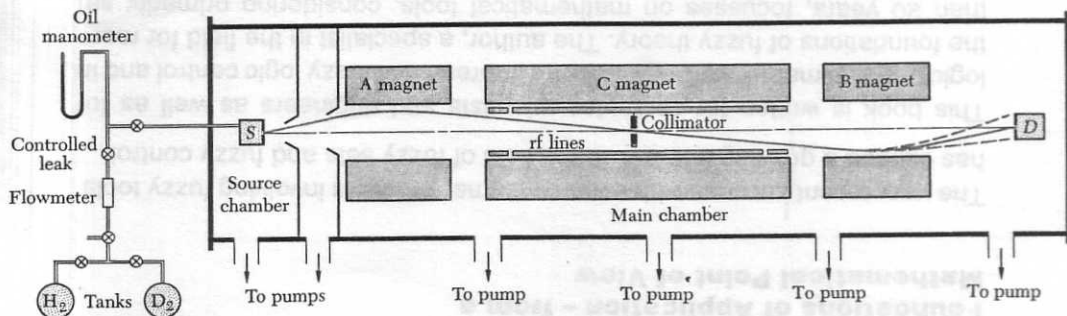


Fig. 6.14 Schematic diagram showing the essential ingredients of the molecular-beam apparatus of the preceding photograph. The source of molecules is the container S, while D denotes some kind of device detecting the molecules arriving at the other end of the apparatus. The inhomogeneous magnetic fields produced

by the A and B magnets exert forces on the very small magnetic moments of the molecules and thus deflect them along the indicated paths. The experiments detect the effect of radio-frequency radiation acting on the molecules in the experimental region of the C magnet.

and closed. When the disks are properly aligned and not rotating, all molecules can reach the detector by passing through the slots in both disks. But when the disks are rotating, molecules passing through the slot in the first disk can only reach the detector if their velocity is such that the time of flight required for them to travel to the second disk is equal to the time required for one revolution of the disks (or an integral number of such revolutions).[†] Otherwise they strike the solid part of the second disk and are stopped. Hence different angular velocities of rotation of the disks allow molecules of different speeds to reach the detector. Measurements of the relative number of molecules arriving at the detector per second thus allow a direct check of the molecular velocity distribution. The validity of the Maxwellian distribution has been well confirmed by such experiments.

The phenomenon of effusion also has several practical applications outside the realm of molecular beams. Returning to the relation (39), note that a knowledge of the absolute temperature T and mean pressure $\bar{p}$ of the gas allows us to calculate n and $\bar{v}$. Thus the equation of state of an ideal gas gives $n = \bar{p}/kT$. Furthermore, the mean speed $\bar{v}$ of a molecule is approximately equal to its most probable speed (32); thus $\bar{v} \propto (kT/m)^{1/2}$. Hence (39) yields

$$\mathcal{F}_0 \propto \frac{\bar{p}}{\sqrt{mT}}. \quad (42)$$

The rate of effusion $\mathcal{F}_0$ is seen to depend on the mass of the molecule because a lighter molecule has a higher mean speed than a heavier one and thus effuses more rapidly. This property can be exploited to achieve a practical method for the separation of isotopes. Suppose that a container is closed off by a membrane which has very many small holes through which molecules can effuse. If this container is surrounded by a vacuum on the outside and is filled with a gas mixture of two isotopes at some initial time, then the relative concentration of the isotope of larger molecular weight will increase in the container as time goes on. Similarly, the gas pumped off from the surrounding vacuum will be more concentrated in the lighter isotope. This method of isotope separation has been of practical importance in obtaining uranium rich in ^{235}U , the isotope which undergoes ready nuclear fission and is thus of great importance for the operation of nuclear power reactors (or the manufacture of nuclear weapons). Ordinary uranium consists mostly of the isotope ^{238}U , but by using the chemical compound uranium hexafluoride (UF_6), which is a gas at room tempera-

[†] Experimentally one can readily distinguish between observations corresponding to such different integral numbers.

ture, one can separate by effusion the slightly lighter $^{235}\text{UF}_6$ molecules from the much more abundant and slightly heavier $^{238}\text{UF}_6$ molecules. Since the difference in mass between these molecules is so small, the process of effusion must be repeated many times in succession to achieve an appreciably increased concentration of the ^{235}U isotope.

6.5 The Equipartition Theorem

The canonical distribution in its classical form (10) is a function of coordinates and momenta which are continuous variables. Any calculation of mean values can thus be reduced to a calculation of integrals rather than of discrete sums. Under certain conditions the mean energy of a system can then be computed in a particularly simple way.

To be specific, consider any system described classically in terms of f coordinates $q_1, \dots, q_f$ and f corresponding momenta $p_1, \dots, p_f$. Its energy E is then a function of these variables, i.e., $E = E(q_1, \dots, p_f)$. It is frequently the case that this energy is of the form

$$E = \epsilon_i(p_i) + E'(q_1, \dots, p_f) \quad (43)$$

where ϵ_i is a function of the particular momentum p_i only and where E' may depend on all coordinates and momenta *except* p_i . (For example, the functional form (43) may arise because the kinetic energy of a particle depends only on its momentum components while its potential energy depends only on its position.) Suppose that the system under consideration is in thermal equilibrium with a heat reservoir at an absolute temperature T . What then is the mean value of the energy contribution ϵ_i in (43)?

The probability of finding the system with its coordinates and momenta in a range near $\{q_1, \dots, q_f, p_1, \dots, p_f\}$ is given by the canonical distribution (10) with the constant C determined by (11). By its definition, the mean value of ϵ_i is found by performing the appropriate sum (or integral) over all possible states of the system, i.e.,

$$\bar{\epsilon}_i = \frac{\int e^{-\beta E(q_1, \dots, p_f)} \epsilon_i dq_1 \dots dp_f}{\int e^{-\beta E(q_1, \dots, p_f)} dq_1 \dots dp_f} \quad (44)$$

where the integrals extend over all possible values of all the coordinates $q_1, \dots, q_f$ and all the momenta $p_1, \dots, p_f$. By virtue of (43), the expression (44) becomes

$$\begin{aligned} \bar{\epsilon}_i &= \frac{\int e^{-\beta(\epsilon_i + E')} \epsilon_i dq_1 \dots dp_f}{\int e^{-\beta(\epsilon_i + E')} dq_1 \dots dp_f} \\ &= \frac{\int e^{-\beta \epsilon_i} \epsilon_i dp_i \int e^{-\beta E'} dq_1 \dots dp_f}{\int e^{-\beta \epsilon_i} dp_i \int e^{-\beta E'} dq_1 \dots dp_f} \end{aligned}$$

where we have used the multiplicative property of the exponential function. The primes on the last integral signs indicate that these integrals extend over all the coordinate q and momenta p except p_i . But since the primed integral in the numerator is the same as that in the denominator, the resulting cancellation leads immediately to the simple result

$$\bar{\epsilon}_i = \frac{\int e^{-\beta \epsilon_i} \epsilon_i dp_i}{\int e^{-\beta \epsilon_i} dp_i}. \quad (45)$$

In short, since ϵ_i involves only the variable p_i , all the other variables are irrelevant in calculating the mean value of ϵ_i .

The expression (45) can be simplified further by relating the integral in the numerator to that in the denominator. Thus

$$\bar{\epsilon}_i = \frac{-\frac{\partial}{\partial \beta} \left(\int e^{-\beta \epsilon_i} dp_i \right)}{\int e^{-\beta \epsilon_i} dp_i}$$

$$\text{or} \quad \bar{\epsilon}_i = -\frac{\partial}{\partial \beta} \ln \left(\int_{-\infty}^{\infty} e^{-\beta \epsilon_i} dp_i \right) \quad (46)$$

where the explicit limits in the integral reflect the fact that the momentum p_i can assume all possible values from $-\infty$ to $+\infty$.

Consider now the specific case where ϵ_i is a quadratic function of p_i , as it would be if it represents a kinetic energy. In short, suppose that ϵ_i is of the form

$$\epsilon_i = b p_i^2 \quad (47)$$

where b is some constant. Then the integral in (46) becomes

$$\int_{-\infty}^{\infty} e^{-\beta \epsilon_i} dp_i = \int_{-\infty}^{\infty} e^{-\beta b p_i^2} dp_i = \beta^{-(1/2)} \int_{-\infty}^{\infty} e^{-b y^2} dy$$

where we have introduced the variable $y \equiv \beta^{1/2} p_i$. Hence

$$\ln \left(\int_{-\infty}^{\infty} e^{-\beta \epsilon_i} dp_i \right) = -\frac{1}{2} \ln \beta + \ln \left(\int_{-\infty}^{\infty} e^{-b y^2} dy \right).$$

But the integral on the right does *not* involve β at all. The differentiation of (46) yields thus simply

$$\bar{\epsilon}_i = -\frac{\partial}{\partial \beta} \left(-\frac{1}{2} \ln \beta \right) = \frac{1}{2\beta}$$

or

$$\boxed{\bar{\epsilon}_i = \frac{1}{2} kT.} \quad (48)$$

Note that, although the starting point (44) of our calculation involved

a formidable array of integrals, we were able to derive the final result (48) without evaluating a single one.

If the functional forms (43) and (47) had been the same except for involving a coordinate q_i instead of a momentum p_i , all our previous arguments would be identical and would thus lead again to (48).

Hence we have established the following general statement, known as the *equipartition theorem*:

If a system described by classical statistical mechanics is in equilibrium at the absolute temperature T , every independent quadratic term in its energy has a mean value equal to $\frac{1}{2}kT$. (48a)

6.6 Applications of the Equipartition Theorem

Specific heat of a monatomic ideal gas

The energy of a molecule in such a gas is simply its kinetic energy (12), i.e.,

$$\epsilon = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2). \quad (49)$$

By virtue of the equipartition theorem, the mean value of each of the three terms in this expression is equal to $\frac{1}{2}kT$. Hence it follows immediately that

$$\bar{\epsilon} = \frac{3}{2}kT. \quad (50)$$

Since one mole of gas consists of Avogadro's number N_a of molecules, the mean energy of the gas is thus,

$$\text{per mole,} \quad \bar{E} = N_a(\frac{3}{2}kT) = \frac{3}{2}RT \quad (51)$$

where $R \equiv N_a k$ is the gas constant. By virtue of (5.23), the molar specific heat c_V at constant volume is then equal to

$$c_V = \left(\frac{\partial \bar{E}}{\partial T} \right)_V = \frac{3}{2}R. \quad (52)$$

This result agrees with that previously obtained in (5.26) on the basis of quantum-mechanical reasoning applied to a gas sufficiently dilute to be ideal and nondegenerate.†

† In accordance with (37), quantum effects should indeed be unimportant for a sufficiently dilute gas. The agreement between the classical and quantum results is thus to be expected.

2. Kinetic energy of a molecule in any gas

Consider any gas, not necessarily ideal. The energy of any molecule of mass m can then be written in the form

$$\epsilon = \epsilon^{(k)} + \epsilon', \quad \text{where } \epsilon^{(k)} = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2).$$

- 1 The first term $\epsilon^{(k)}$ is the kinetic energy of the molecule and depends on the momentum components p_x, p_y, p_z , of its center of mass. The term
- 2 ϵ' may involve the position of the center of mass of the molecule (if the molecule is located in an external force field or if it interacts appreciably with other molecules); it may involve coordinates and momenta describing the rotation or vibration of the atoms of the molecule with respect to its center of mass (if the molecule is not monatomic); but it does not involve the center-of-mass momentum $\mathbf{p}$. Hence the equipartition theorem allows us again to conclude immediately that

$$\overline{\frac{1}{2m} p_x^2} = \overline{\frac{1}{2} m v_x^2} = \frac{1}{2} kT \quad (53)$$

or

$$\overline{v_x^2} = \frac{kT}{m}. \quad (54)$$

- 1 Since $\overline{v_x} = 0$ by symmetry, as already mentioned in (26), the result (54) represents also the dispersion $\overline{(\Delta v_x)^2}$ of the velocity component v_x .
- 2 The three quadratic terms of the kinetic energy $\epsilon^{(k)}$ then yield for its mean value, as in (50), the result

$$\overline{\epsilon^{(k)}} = \frac{3}{2} kT. \quad (55)$$

3. Brownian motion

- 1 Consider a macroscopic particle of mass m (about a micron in size) suspended in a fluid at an absolute temperature T . The energy of this particle can again be written in the form

$$\epsilon = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + \epsilon'.$$

- 2 Here the first term is the kinetic energy involving the velocity $\mathbf{v}$ or momentum $\mathbf{p} = m\mathbf{v}$ of the center-of-mass motion of the particle, while
- 3 ϵ' is the energy associated with the motion of all the atoms of the particle with respect to its center of mass. The equipartition theorem leads then again to the results (53) and (54) so that

$$\overline{v_x^2} = \frac{kT}{m}. \quad (56)$$

Since the mean value $\bar{v}_x = 0$ vanishes by symmetry, (56) gives directly the dispersion of the velocity component v_x . The result (56) thus shows immediately that the particle does not simply remain at rest, but must always exhibit a fluctuating velocity. The existence of the phenomenon of Brownian motion, discussed in Sec. 1.4, is thus an immediate consequence of our theory. The quantitative statement (56) also shows explicitly that, if the mass m of the particle is sufficiently large, the fluctuations become so small as to be unobservable.

4. Harmonic oscillator

Consider a particle of mass m performing simple harmonic oscillations in one dimension. Its energy is then given by

$$\epsilon = \frac{1}{2m} p_x^2 + \frac{1}{2} \alpha x^2. \quad (57)$$

Here the first term is the kinetic energy of the particle whose momentum is denoted by p_x . The second term is its potential energy if a displacement x of the particle produces a restoring force $-\alpha x$, where α is a constant (called the *spring constant*). Suppose that the oscillator is in equilibrium with a heat reservoir at a temperature T high enough so that the oscillator can be described in terms of classical mechanics. Then the equipartition theorem (48) can be immediately applied to each of the quadratic terms in (57) to give for the mean energy of the oscillator

$$\bar{\epsilon} = \frac{1}{2}kT + \frac{1}{2}kT = kT. \quad (58)$$

6.7 The Specific Heat of Solids

As a last application of the equipartition theorem, we shall discuss the specific heat of solids at temperatures sufficiently high so that a classical description is valid. Consider thus any simple solid consisting of N atoms; e.g., the solid might be copper, gold, aluminum, or diamond. By virtue of the mutual forces between neighboring atoms, the situation of stable mechanical equilibrium of the solid is that in which its atoms are located at regular positions in a crystal lattice. Each atom is, however, free to move by small amounts about its equilibrium position. The force (due to neighboring atoms) tending to restore the atom to its equilibrium position vanishes, of course, when the atom is actually located at its equilibrium position. Since the displacement of the atom

- 4 from its equilibrium position is always quite small, the restoring force must, in first approximation, be simply proportional to the atomic displacement. This approximation, which is usually excellent, thus implies that the atom performs about its equilibrium position simple harmonic motion in three dimensions.

4. With proper choice of orientation of the x, y, z coordinate axes, the motion of an atom along any one of these axes, say along the x direction, is then a simple harmonic motion with an associated energy ϵ_x of the form (57), i.e.,

$$\epsilon_x = \frac{1}{2m} p_x^2 + \frac{1}{2} \alpha x^2. \quad (59)$$

- 1 Here p_x denotes the x component of momentum of the atom while x denotes the x component of displacement of the atom from its equilibrium position. In writing (59), we have supposed that the atom has a mass m and is subject to a restoring force of spring constant α . Correspondingly, the (angular) frequency of oscillation of the atom along the x direction is then given by

$$\omega = \sqrt{\frac{\alpha}{m}}. \quad (60)$$

2. Similar expressions hold for the energies ϵ_y and ϵ_z associated with the motion of the atom in the y and z directions. The total energy of the atom is thus of the form

$$\epsilon = \epsilon_x + \epsilon_y + \epsilon_z. \quad (61)$$

- 2 If the solid is in equilibrium at an absolute temperature T which is high enough so that the approximation of classical statistical mechanics is valid, the equipartition theorem is immediately applicable to each of the quadratic terms of (59). The mean value of ϵ_x is thus simply

$$\bar{\epsilon}_x = \frac{1}{2}kT + \frac{1}{2}kT = kT. \quad (62)$$

Similarly $\bar{\epsilon}_y = \bar{\epsilon}_z = kT$. By (61), the mean energy of an atom is then

$$\bar{\epsilon} = 3kT.$$

- 1 The mean energy of one mole of the solid, containing Avogadro's number N_a of atoms, is thus simply

$$E = 3N_a kT = 3RT \quad (63)$$

- 2 where $R = N_a k$ is the gas constant. In accordance with (5.23), the molar specific heat c_V at constant volume of the solid is then given by

$$c_V = \left(\frac{\partial E}{\partial T} \right)_V$$

or

$$c_V = 3R. \quad (64)$$

Using the numerical value (5.4) of R , this becomes†

$$c_V \approx 25 \text{ joules mole}^{-1} \text{ deg}^{-1}. \quad (65)$$

Note the extreme generality of the result (64). It is completely independent of the atomic mass or of the magnitude of the spring constant α . If the solid contained atoms having different masses or having different spring constants, the result (64) would thus still remain valid. If the solid were not isotropic, the restoring force on an atom would have a different magnitude along different directions and the spring constants for the x , y , and z directions would also differ correspondingly; but the value of the mean energy per atom would still be $3kT$ and (64) would still be valid. Indeed, the rigorous analysis of the simultaneous vibration of all the atoms in the solid shows that a description in terms of individual motions of single atoms is inadequate and that simple harmonic motion is really performed by atoms moving together in groups of different sizes.‡ But since (64) is independent of the masses and spring constants, it still remains valid. The only limitation on (64) is thus that the temperature be sufficiently high so that the classical approximation is valid. Hence (64) asserts that:

At sufficiently high temperatures, all solids have the same temperature-independent molar specific heat c_V equal to $3R$. (66)

We shall see presently that in the case of most solids (diamond being a conspicuous exception) room temperature is sufficiently high for the approximate validity of the classical discussion.

Historically, the validity of statement (66) was first discovered empirically and is known as the law of Dulong and Petit. Table 6.1 lists directly measured values of the molar specific heat c_p (at constant pressure) for some solids at room temperature. The values of the molar specific heat c_V (at constant volume) can be obtained from the corre-

Solid	c_p	c_V
Aluminum	24.4	23.4
Bismuth	25.6	25.3
Cadmium	26.0	24.6
Carbon (diamond)	6.1	6.1
Copper	24.5	23.8
Germanium	23.4	23.3
Gold	25.4	24.5
Lead	26.8	24.8
Platinum	25.9	25.4
Silicon	19.8	19.8
Silver	25.5	24.4
Sodium	28.2	25.6
Tin (metallic)	26.4	25.4
Tungsten	24.4	24.4

Table 6.1 Values of the molar specific heat c_p at constant pressure and the molar specific heat c_V at constant volume for some simple solids at a temperature $T = 298^\circ\text{K}$. The values of c_V were deduced from the directly measured values of c_p by applying some small corrections. All values are expressed in units of joules $\text{deg}^{-1} \text{ mole}^{-1}$. [Data taken from Dwight E. Gray (ed.), *American Institute of Physics Handbook*, 2d ed., p. 4-48 (McGraw-Hill Book Company, New York, 1963).]

† In terms of calories, $c_V \approx 6 \text{ cal mole}^{-1} \text{ deg}^{-1}$.

‡ That is, by the normal modes of the solid.

sponding values of c_p after applying some small corrections.[†] It is seen that the values of c_v listed in the table are, on the whole, in rather good agreement with the value (65) predicted by the classical theory. Bad discrepancies occur, however, in the cases of silicon and, particularly, diamond. The reason is that quantum effects are still important for these solids even at a temperature as high as 300°K.

Validity of the classical approximation

Let us then ask under what circumstances the preceding classical discussion is expected to be valid. The criterion is again provided by the condition (3). Consider thus a vibrating atom which has an energy ϵ_x associated with its motion in the x direction. By virtue of the equipartition theorem applied to (59), the momentum p_x of the atom is then such that

$$\frac{1}{2m} \overline{p_x^2} = \frac{1}{2} kT.$$

The momentum of an atom thus has a typical magnitude p_0 of the order of

$$p_0 \approx \sqrt{\overline{p_x^2}} \approx \sqrt{mkT}. \quad (67)$$

To assure the validity of the classical description, quantum effects should not prevent us from localizing an atom within a typical distance s_0 of the order of the mean amplitude of vibration of this atom. But the equipartition theorem applied to (59) yields

$$\frac{1}{2} \alpha \overline{x^2} = \frac{1}{2} kT.$$

The typical magnitude s_0 of the atomic displacement is thus of the order of

$$s_0 \approx \sqrt{\overline{x^2}} \approx \sqrt{\frac{kT}{\alpha}}. \quad (68)$$

Hence the condition (3) that the Heisenberg uncertainty principle be of negligible significance becomes simply

[†] It is quite easy to make specific heat measurements on solids at constant atmospheric pressure, allowing the volume of the solid to change slightly, but very difficult to devise an experimental arrangement guaranteed to prevent the solid from expanding its volume when its temperature is raised. Since the volume of a solid changes only slightly as a result of a temperature change, the difference between c_p and c_v is, however, quite small. It can readily be calculated from a knowledge of the measured macroscopic properties of the particular solid under consideration.

$$s_0 p_0 \approx kT \sqrt{\frac{m}{\alpha}} \gg \hbar$$

or

$$kT \gg \hbar \omega \quad (69)$$

where ω is, by (60), the typical (angular) frequency of oscillation of an atom in the solid. Equivalently the criterion (69) for the validity of the classical approximation can be written in the form

$$T \gg \Theta, \quad \text{where } \Theta \equiv \frac{\hbar \omega}{k} \quad (70)$$

is a temperature parameter characteristic of the solid under consideration.

3. Numerical estimates

The magnitude of the atomic frequency of oscillation ω can be estimated from the elastic properties of the solid under consideration. Suppose, for example, that a small pressure Δp is applied to a solid†; as a result, the volume V of the solid will decrease by some small amount

1. ΔV . The quantity κ defined by

$$\kappa \equiv -\frac{1}{V} \frac{\Delta V}{\Delta p} \quad (71)$$

is called the *compressibility* of the solid (the minus sign being introduced to make κ positive). It is readily measured and provides some information about the forces between the atoms in the solid.

Let us then try to deduce from the compressibility κ a rough estimate of the net force F acting on an atom when it is displaced from its equilibrium position in the solid. Imagine, for simplicity, that the atoms of the solid are located at the centers of cubes of edge length a so that the interatomic spacing is also equal to a . An excess pressure Δp applied to a surface of the solid corresponds then to a force $F = a^2 \Delta p$ applied to the area a^2 occupied by a single atom (see Fig. 6.15). Furthermore, the fractional volume

change of the solid under the influence of the excess pressure Δp is equal to the fractional volume change occupied per atom so that

$$\frac{\Delta V}{V} = \frac{\Delta(a^3)}{a^3} = \frac{3a^2 \Delta a}{a^3} = \frac{3\Delta a}{a}.$$

Using the definition (71) of the compressibility, the force F on an atom is then related to Δa by

$$F = a^2 \Delta p = a^2 \left(-\frac{1}{\kappa} \frac{\Delta V}{V} \right) = -\frac{a^2}{\kappa} \frac{3\Delta a}{a}$$

or

$$F = -\alpha \Delta a$$

where the constant α , relating the force F to the displacement Δa of the atom from its equilibrium position, is given by

$$\alpha = \frac{3a}{\kappa}. \quad (72)$$

With our simple approximation of an assumed cubic lattice of atoms in the solid, the estimated frequency of vibration (60) of an atom in the solid should then be approximately

$$\omega = \sqrt{\frac{\alpha}{m}} \approx \sqrt{\frac{3a}{\kappa m}}. \quad (73)$$

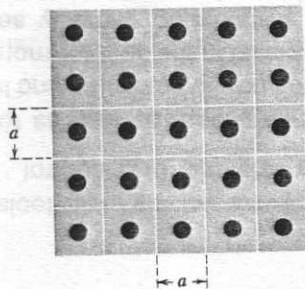


Fig. 6.15 Surface-view of a solid whose atoms are arranged in a simple cubic lattice.

† No confusion should arise between the symbol p used for pressure and the same symbol used elsewhere to denote momentum.

To get an appreciation for the relevant magnitudes, let us estimate ω in the case of copper. The measured parameters of this metal are†:

atomic weight

$$\mu = 63.5,$$

density

$$\rho = 8.95 \text{ gm cm}^{-3},$$

compressibility

$$\kappa = 7.3 \times 10^{-13} \text{ cm}^2 \text{ dyne}^{-1}.$$

From these numbers we find for the atomic mass

$$m = \frac{\mu}{N_a} = \frac{63.5}{6.02 \times 10^{23}} \\ = 1.05 \times 10^{-22} \text{ gm}.$$

Since $\rho = m/a^3$, the interatomic distance is equal to

$$a = \left(\frac{m}{\rho}\right)^{1/3} = \left(\frac{1.05 \times 10^{-22}}{8.95}\right)^{1/3} \\ = 2.34 \times 10^{-8} \text{ cm}.$$

Hence (73) gives for the angular frequency of vibration

$$\omega \approx \left[\frac{3(2.34 \times 10^{-8})}{(7.3 \times 10^{-13})(1.05 \times 10^{-22})} \right]^{1/2} \\ = 3.02 \times 10^{13} \text{ radians/sec},$$

or for the corresponding frequency

$$\nu = \frac{\omega}{2\pi} \approx 4.8 \times 10^{12} \text{ cycles/sec.} \quad (74)$$

This is a frequency lying in the infrared region of the electromagnetic spectrum.

The characteristic temperature Θ defined in (70) is then equal to

$$\Theta \equiv \frac{\hbar\omega}{k} = \frac{(1.054 \times 10^{-27})(3.02 \times 10^{13})}{(1.38 \times 10^{-16})} \\ \approx 230^\circ \text{K.} \quad (75)$$

Hence the classical result $c_V = 3R$ should be valid for copper if $T \gg 230^\circ \text{K}$, i.e., it should begin to be reasonably accurate for temperatures of the order of room temperature or above.

On the other hand, let us consider a solid such as diamond. The atomic mass of one of its carbon atoms is 12, i.e., about 5 times less than that of a copper atom. In addition, diamond is a very hard solid so that its compressibility is very low, about 3 times less than that of copper ($\kappa = 2.26 \times 10^{-13} \text{ cm}^2 \text{ dyne}^{-1}$). Thus the frequency of vibration ω of a carbon atom in diamond is, by (73), much higher than that of a copper atom in copper metal. More precisely, for diamond (density $\rho = 3.52 \text{ gm cm}^{-3}$) the estimated temperature parameter $\Theta \approx 830^\circ \text{K}$. The classical approximation is, therefore, not expected to be applicable to diamond at room temperature and the low value of c_V for diamond in Table 6.1 is by no means surprising.

It is clear that the classical result $c_V = 3R$ must break down at low temperatures where the condition (69) is not satisfied. Indeed, the very general result (5.32) implies that, as the temperature is reduced below the range of validity of (69), the specific heat c_V must ultimately decrease so as to approach zero as $T \rightarrow 0$. Any correct quantum-mechanical calculation must yield this limiting result. If every atom in the solid is assumed to vibrate with the same frequency ω , the quantum-mechanical calculation of the specific heat c_V can be carried out quite easily to give an approximate expression for c_V valid for all temperatures. Details will be left as an exercise in Prob. 6.21.

† The data are taken from Dwight E. Gray (ed.), *American Institute of Physics Handbook*, 2d ed. (McGraw-Hill Book Company, New York, 1963).

Mon. 20-7-1992 } during 2 months
20.14 GMT }
Sat 5-9-1992 } 24 hours in 2-days
2.11.7 GMT }

Summary of Definitions

phase space A cartesian multidimensional space whose axes are labeled by all the coordinates and momenta describing a system in classical mechanics. A point in this space specifies all the coordinates and momenta of the system.

Maxwell velocity distribution The expression

$$f(\mathbf{v}) d^3\mathbf{v} \propto e^{-(1/2)\beta m v^2} d^3\mathbf{v}$$

giving the mean number of molecules having a velocity between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$ in a gas at the absolute temperature T . It is merely a special case of the canonical distribution.

effusion The outflow of molecules from a container through a small hole whose size is much smaller than the molecular mean free path.

Important Relations

If a system described classically is in equilibrium at the absolute temperature T , every independent quadratic term ϵ_i of its energy has a mean value

$$\bar{\epsilon}_i = \frac{1}{2}kT. \quad (i)$$

Suggestions for Supplementary Reading

- O. R. Frisch, "Molecular Beams," *Sci. American* **212**, 58 (May 1965). A good discussion of the wide range of fundamental physical experiments made possible by molecular beam investigations.
- F. Reif, *Fundamentals of Statistical and Thermal Physics*, chap. 7 (McGraw-Hill Book Company, New York, 1965). A somewhat more detailed discussion of the topics of the present chapter.
- F. W. Sears, *An Introduction to Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics*, 2d ed., chaps. 11 and 12 (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1953).
- D. K. C. MacDonald, *Faraday, Maxwell, and Kelvin* (Anchor Books, Doubleday & Company, Inc., Garden City, N.Y., 1964). This book contains a short account of Maxwell's life and scientific work.

Problems

6.1 Phase space of a classical harmonic oscillator

The energy of a one-dimensional harmonic oscillator, whose position coordinate is x and whose momentum is p , is given by

$$E = \frac{1}{2m} p^2 + \frac{1}{2} \alpha x^2$$

where the first term on the right is its kinetic and the second term its potential energy. Here m denotes the mass of the oscillating particle and α the spring constant of the restoring force acting on the particle.

Consider an ensemble of such oscillators, the energy of each oscillator being known to lie between E and $E + \delta E$. Treating the situation classically, indicate in the two-dimensional xp phase space the region of states accessible to the oscillator.

6.2 Ideal gas in a gravitational field

An ideal gas at the absolute temperature T is in equilibrium in the presence of a gravitational field described by an acceleration g in the downward (or $-z$) direction. The mass of each molecule is m .

(a) Use the canonical distribution in its classical form to find the probability $\mathcal{P}(\mathbf{r}, \mathbf{p}) d^3\mathbf{r} d^3\mathbf{p}$ that a molecule has a position between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ and a momentum between $\mathbf{p}$ and $\mathbf{p} + d\mathbf{p}$.

(b) Find (to within a trivial constant of proportionality) the probability $\mathcal{P}'(\mathbf{v}) d^3\mathbf{v}$ that a molecule has a velocity between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$, irrespective of its position in space. Compare this result with the corresponding probability in the absence of a gravitational field.

(c) Find (to within a trivial constant of proportionality) the probability $\mathcal{P}''(z) dz$ that a molecule is located at a height between z and $z + dz$, irrespective of its velocity or its location in any horizontal plane.

6.3 Macroscopic discussion of an ideal gas in a gravitational field

Consider the ideal gas of the last problem from a completely macroscopic point of view. By writing down the condition of mechanical equilibrium for a slice of the gas located between the heights z and $z + dz$, and by using the equation of state (4.92), derive an expression for $n(z)$, the number of molecules per unit volume at a height z . Compare this with the result for $\mathcal{P}''(z) dz$ derived in the last problem on the basis of statistical mechanics.

6.4 Spatial distribution of electrons in a cylindrical electric field

A wire of radius r_0 is coincident with the axis of a metal cylinder of radius R and length L . The wire is maintained at a positive potential of V statvolts with respect to the cylinder. The whole system is at some high absolute temperature T . As a result, electrons emitted from the hot metals form a dilute gas filling the cylindrical container and in equilibrium with it. The density of these electrons is so low that their mutual electrostatic interaction can be neglected.

(a) Use Gauss's theorem to obtain an expression for the electrostatic field which exists at points at a radial distance r from the wire ($r_0 < r < R$). The cylinder of length L may be assumed to be very long so that end effects are negligible.

(b) In thermal equilibrium, the electrons form a gas of variable density which fills the entire space between the wire and cylinder. Using the result of part (a), find how the number n of electrons per unit volume depends on the radial distance r .

(c) Give an approximate criterion determining how low the temperature T must be so that the electron density is small enough to justify the approximation of neglecting the mutual electrostatic interaction between the electrons.

6.5 Determination of large molecular weights by the ultracentrifuge

Consider a macromolecule (i.e., a very large molecule with a molecular weight of several millions) immersed in an incompressible fluid of density ρ at the absolute temperature T . The volume v occupied by one such molecule can be considered known since the volume occupied by a mole of macromolecules can be determined by volume measurements on a solution of macromolecules. A dilute solution of this type is now placed in an ultracentrifuge rotating with a high angular velocity ω . In the frame of reference rotating with the centrifuge, any particle of mass m at rest with respect to this frame is then acted upon by an outward centrifugal force $m\omega^2 r$, where r denotes the distance of the particle from the axis of rotation.

(a) What is the net force acting in this frame of reference on a macromolecule of mass m , if the buoyancy effect of the surrounding fluid is taken into account?

(b) Suppose that equilibrium has been attained in this frame of reference so that the mean number $n(r) dr$ (per unit volume) of macromolecules located at a distance from the axis of rotation between r and $r + dr$ is independent of time. Apply the canonical distribution to find (to within a constant of proportionality) the number $n(r) dr$ as a function of r .

(c) Measurements of the relative number $n(r)$ of molecules as a function of r can be made by measuring the absorption of light by the solution. Show how such measurements can be used to deduce the mass m of a macromolecule.

***6.6 Spatial separation of magnetic atoms in an inhomogeneous magnetic field**

An aqueous solution at room temperature T contains a small concentration of magnetic atoms, each of which has a spin $\frac{1}{2}$ and magnetic moment μ_0 . The solution is placed in an external magnetic field. The magnitude of this field is inhomogeneous over the volume of the solution. To be specific, the z component B of this magnetic field is a uniformly increasing function of z , assuming a value B_1 at the bottom of the solution where $z = z_1$ and a larger value B_2 at the top of the solution where $z = z_2$.

(a) Let $n_+(z) dz$ denote the mean number of magnetic atoms whose magnetic moment points along the z direction and which are located between z and $z + dz$. What is the ratio $n_+(z_2)/n_+(z_1)$?

(b) Let $n(z) dz$ denote the total mean number of magnetic atoms (of both directions of spin orientation) located between z and $z + dz$. What is the ratio $n(z_2)/n(z_1)$? Is it less than, equal to, or greater than unity?

(c) Make use of the fact that $\mu_0 B \ll kT$ to simplify the answers to the preceding questions.

(d) Estimate numerically the magnitude of the ratio $n(z_2)/n(z_1)$ at room temperature if $\mu_0 \approx 10^{-20}$ erg/gauss is of the order of a Bohr magneton, $B_1 = 0$, and $B_2 = 5 \times 10^4$ gauss.

6.7 Most probable energy of a molecule in a gas

What is the most probable kinetic energy $\bar{\epsilon}$ of a molecule described by a Maxwellian velocity distribution? Is it equal to $\frac{1}{2}m\bar{v}^2$, where $\bar{v}$ is the most probable speed of the molecule?

6.8 Temperature dependence of effusion

Molecules of a gas enclosed within a container effuse through a small hole into a surrounding vacuum. Suppose that the absolute temperature of the gas in the container is doubled while its pressure is kept constant.

(a) By what factor does the number of molecules escaping per second through the hole change?

(b) By what factor does the force change which is exerted on a vane suspended at some distance in front of the hole?

6.9 Mean kinetic energy of an effusing molecule

The molecules of a monatomic ideal gas are escaping by effusion through a small hole in a wall of an enclosure maintained at an absolute temperature T . By physical reasoning (without actual calculation) do you expect the mean kinetic energy $\bar{\epsilon}_0$ of a molecule in the effusing beam to be equal to, greater than, or less than the mean kinetic energy $\bar{\epsilon}_i$ of a molecule inside the enclosure?

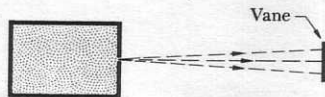


Fig. 6.16 An effusing beam impinging upon a vane.

6.10 Pressure drop in gas container having a small leak

A thin-walled vessel of volume V , kept at constant temperature T , contains a gas which slowly leaks out through a small hole of area A . The outside pressure is low enough so that leakage back into the vessel is negligible. Estimate the time required for the pressure in the vessel to decrease to half of its original value. Express your answer in terms of A , V , and the mean molecular speed $\bar{v}$.

6.11 Cryogenic (i.e., low-temperature) pumping

Gases can be removed from a container by lowering the temperature of some of its walls. This is a method commonly used to achieve the good vacuums necessary for many physical experiments. To illustrate the principle of the method, consider a spherical bulb 10 cm in radius which is maintained at room temperature (300°K) except for a 1-cm² spot which is kept at liquid nitrogen temperature (77°K). The bulb contains water vapor originally at a pressure of 0.1 mm of mercury. Assuming that every water molecule striking the cold area condenses and sticks to the surface, estimate the time required for the pressure to decrease to 10⁻⁶ mm of mercury.

6.12 Separation of isotopes by effusion

A vessel has porous walls containing very many tiny holes. Gas molecules can pass through these holes by effusion and then be pumped off to some collecting chamber. The vessel is filled with a dilute gas consisting of two kinds of molecules which have different masses m_1 and m_2 by virtue of the fact that they contain two different isotopes of the same atom. Let us denote by c_1 the concentration of the first type of molecules in the vessel, and by c_2 the concentration of the second type. (The concentration c_i is the ratio of the number of molecules of type i to the total number of molecules.) These concentrations can be kept constant in the vessel by providing a steady flow of fresh gas through it so as to replenish any gas that has effused.

(a) Let c_1' and c_2' denote the concentrations of the two types of molecules in the collecting chamber. What is the ratio c_2'/c_1' ?

(b) By using the gas UF_6 , one can attempt to separate ^{235}U from ^{238}U , the first of these isotopes being the one useful in the initiation of nuclear-fission reactions. The molecules in the vessel are then $^{238}\text{U}^{19}\text{F}_6$ and $^{235}\text{U}^{19}\text{F}_6$. (The concentrations of these molecules, corresponding to the natural abundance of the two uranium isotopes, are $c_{238} = 99.3$ per cent and $c_{235} = 0.7$ per cent.) Calculate the corresponding ratio c'_{235}/c'_{238} of the molecules collected after effusion. Express the result in terms of the original concentration ratio c_{235}/c_{238} .

6.13 Change of concentration as a result of effusion

A container has as one of its walls a membrane containing many small holes. If the container is filled with gas at some moderate pressure $\bar{p}$, this gas will escape by effusion into the vacuum surrounding the container. It is found

that when the container is filled with helium gas at room temperature and at a pressure $\bar{p}$, the pressure will have fallen to $\frac{1}{2}\bar{p}$ after one hour.

Suppose that the container is filled at room temperature and at a total pressure $\bar{p}$ with a mixture of helium (He) and neon (Ne), the atomic concentrations of both gases being 50 per cent (i.e., 50 per cent of the atoms are He and 50 per cent of them are Ne). What will be the ratio $n_{\text{Ne}}/n_{\text{He}}$ of the atomic concentrations of Ne to He after one hour? Express your answer in terms of the atomic weights μ_{Ne} of neon and μ_{He} of helium.

6.14 Calculation of mean values for a molecule in a gas

A gas of molecules, each having mass m , is at rest in thermal equilibrium at the absolute temperature T . Denote the velocity of a molecule by $\mathbf{v}$, its three cartesian components of velocity by v_x , v_y , and v_z , and its speed by v . Find the following mean values:

- (a) $\overline{v_x}$,
- (b) $\overline{v_x^2}$,
- (c) $\overline{v^2 v_x}$,
- (d) $\overline{v_x^2 v_y}$,
- (e) $\overline{(v_x + bv_y)^2}$, where b is a constant.

(Suggestion: Symmetry arguments and the equipartition theorem should suffice to answer all these questions without any significant calculation.)

6.15 Doppler broadening of spectral lines

A gas of atoms, each of mass m , is maintained at the absolute temperature T inside an enclosure. The atoms emit light which passes (in the x direction) through a window of the enclosure and can then be observed as a spectral line in a spectroscope. A stationary atom would emit light at the sharply defined frequency ν_0 . But, because of the Doppler effect, the frequency of the light observed from an atom having an x component of velocity v_x is not simply equal to the frequency ν_0 , but is given approximately by

$$\nu = \nu_0 \left(1 + \frac{v_x}{c} \right)$$

where c is the velocity of light. As a result, not all of the light arriving at the spectroscope is at the frequency ν_0 ; instead it is characterized by some intensity distribution $I(\nu) d\nu$ giving the fraction of light intensity lying in the frequency range between ν and $\nu + d\nu$.

- (a) Calculate the mean frequency $\bar{\nu}$ of the light observed in the spectroscope.
- (b) Calculate the dispersion $(\Delta\nu)^2 = (\nu - \bar{\nu})^2$ in the frequency of the light observed in the spectroscope.
- (c) Show how measurements of the width $\Delta\nu \equiv [(\Delta\nu)^2]^{1/2}$ of a spectral line observed in the light coming from a star allow one to determine the temperature of that star.

6.16 Specific heat of an adsorbed mobile monolayer

If the surface of some solid is maintained in a reasonably good vacuum, a single layer of molecules, one molecular diameter thick, can form on this surface. (The molecules are then said to be adsorbed on the surface.) The molecules are held to this surface by forces exerted on them by the atoms of the solid, but they may be quite free to move in two dimensions on this surface. They form then, to good approximation, a classical two-dimensional gas. If the molecules are monatomic and the absolute temperature is T , what is the specific heat per mole of molecules thus adsorbed on a surface of fixed size?

6.17 Temperature dependence of the electrical resistivity of a metal

The electrical resistivity ρ of a metal is proportional to the probability that an electron is scattered by the vibrating atoms in the lattice, and this probability is in turn proportional to the mean square amplitude of vibration of these atoms. How does the electrical resistivity ρ of the metal depend on its absolute temperature in the range near room temperature, or above, where classical statistical mechanics can validly be applied to discuss the vibrations of the atoms in the metal?

6.18 Theoretical limiting accuracy of a weight measurement

A very sensitive spring balance consists of a quartz spring suspended from a fixed support. The spring constant is α , i.e., the restoring force of the spring is $-\alpha x$ if the spring is stretched by an amount x . The balance is at an absolute temperature T in a location where the acceleration due to gravity is g .

(a) If a very small object of mass M is suspended from the spring, what is the mean resultant elongation $\bar{x}$ of the spring?

(b) What is the magnitude $(\Delta x)^2 \equiv (x - \bar{x})^2$ of the thermal fluctuations of the object about its equilibrium position?

(c) It becomes impracticable to measure the mass of an object when the fluctuations are so large that $[(\Delta x)^2]^{1/2} \geq \bar{x}$. What is the minimum mass M which can be measured with this balance?

6.19 Specific heat of anharmonic oscillators

Consider a one-dimensional oscillator (not simple harmonic) which is described by a position coordinate x and by a momentum p and whose energy is given by

$$\epsilon = \frac{p^2}{2m} + bx^4 \quad (i)$$

where the first term on the right is its kinetic energy and the second term is its potential energy. Here m denotes the mass of the oscillator and b is some constant. Suppose that this oscillator is in thermal equilibrium with a heat reservoir at a temperature T high enough so that the approximation of classical mechanics is a good one.

- (a) What is the mean kinetic energy of this oscillator?
 (b) What is its mean potential energy?
 (c) What is its mean total energy?
 (d) Consider an assembly of weakly interacting particles, each vibrating in one dimension so that its energy is given by (i). What is the specific heat at constant volume per mole of these particles?

(Suggestion: There is no need to evaluate explicitly any integral to answer these questions.)

6.20 Specific heat of a highly anisotropic solid

Consider a solid which has a highly anisotropic crystalline layer structure. Each atom in this structure can be regarded as performing simple harmonic oscillations in three dimensions. The restoring forces in directions parallel to a layer are very large; hence the natural frequencies of oscillations in the x and y directions lying within the plane of a layer are both equal to a value $\omega_{\parallel}$ which is so large that $\hbar\omega_{\parallel} \gg 300k$, the thermal energy kT at room temperature. On the other hand, the restoring force perpendicular to a layer is quite small; hence the frequency of oscillation $\omega_{\perp}$ of an atom in the z direction perpendicular to a layer is so small that $\hbar\omega_{\perp} \ll 300k$. On the basis of this model, what is the molar specific heat (at constant volume) of this solid at 300°K ?

6.21 Quantum theory of the specific heat of solids

To treat the atomic vibrations in a solid by quantum mechanics, use as a simplifying approximation a model which assumes that each atom of the solid vibrates independently of the other atoms with the same angular frequency ω in each of its three directions. The solid consisting of N atoms is then equivalent to an assembly of $3N$ independent one-dimensional oscillators vibrating with the frequency ω . The possible quantum states of every such oscillator have discrete energies given by

$$\epsilon_n = (n + \frac{1}{2})\hbar\omega \quad (i)$$

where the quantum number n can assume the possible values $n = 0, 1, 2, 3, \dots$

(a) Suppose that the solid is in equilibrium at the absolute temperature T . By using the energy levels (i) and the canonical distribution, proceed as in Prob. 4.22 to calculate the mean energy $\bar{\epsilon}$ of an oscillator and thus also the total mean energy $E = N\bar{\epsilon}$ of the vibrating atoms in the solid.

(b) Using the result of part (a), proceed as in Prob. 5.20 to calculate the molar specific heat c_V of the solid.

264 Canonical Distribution in the Classical Approximation

(c) Show that the result of part (b) can be expressed in the form

$$c_V = 3R \frac{w^2 e^w}{(e^w - 1)^2} \quad (\text{ii})$$

where

$$w \equiv \frac{\hbar\omega}{kT} = \frac{\Theta}{T} \quad (\text{iii})$$

and where $\Theta \equiv \hbar\omega/k$ is the temperature parameter previously defined in (70).

(d) Show that, when $T \gg \Theta$, the result (ii) approaches properly the classical value $c_V = 3R$.

(e) Show that the expression (ii) for c_V approaches properly the value zero as $T \rightarrow 0$.

(f) Find an approximate expression for the result (ii) in the limit when $T \ll \Theta$.

(g) Make a rough sketch of c_V as a function of the absolute temperature T .

(h) Apply the criterion (1) to find below what temperature the classical approximation is *not* expected to be applicable. Compare your result with the condition (69) for the applicability of the classical theory of specific heats.

[Using the approximations made in this problem, Einstein first derived the expression (ii) in 1907. Using the novel quantum ideas, he was thus able to account for the experimentally observed specific heat behavior which had been inexplicable on the basis of the classical theory.]

Appendix

A.1 Gaussian Distribution

Consider the binomial distribution derived in (2.14),

$$P(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} \quad (1)$$

where $q = 1 - p$. When N is large, the calculation of the probability $P(n)$ appears difficult since it requires the computation of the factorials of large numbers. It then becomes possible, however, to use approximations which allow us to transform the expression (1) into a particularly simple form.

The simplifying feature, already pointed out in Sec. 2.3, is that the probability $P(n)$ tends to exhibit a maximum which is quite pronounced when N is large. The probability $P(n)$ thus becomes negligibly small whenever n differs appreciably from the particular value $n = \bar{n}$ where P is maximum. Hence the region ordinarily of interest, that where the probability $P(n)$ is *not* negligible, consists only of those values of n which do not differ very much from $\bar{n}$. But in this relatively small region an approximate expression for $P(n)$ can readily be found. This expression can then be used for all values of n where the probability P is not negligibly small, i.e., in the entire domain where a knowledge of P is ordinarily of interest.

It is thus sufficient to investigate the behavior of $P(n)$ near the position $\bar{n}$ of its maximum. Note first that, unless $p \approx 0$ or $q \approx 0$, this value $\bar{n}$ is neither very close to 0 nor to N ; thus $\bar{n}$ itself is also a large number when N is large. The numbers n in the region of interest near $\bar{n}$ are thus also large. But when n is large, $P(n)$ changes relatively little when n changes by unity; i.e.,

$$|P(n+1) - P(n)| \ll P(n)$$

so that P is a slowly varying function of n . It is, therefore, a good approximation to regard P as a smooth function of a continuous variable n , although only integral values of n are of physical relevance. A second useful observation concerns the fact that the logarithm of P is a much more slowly varying function of n than P itself. Instead of dealing directly with P , it is thus easier to investigate the behavior of $\ln P$ and to find for $\ln P$ a good approximation valid in a large domain of the variable n .

Taking the logarithm of (1), we thus obtain

$$\ln P = \ln N! - \ln n! - \ln (N - n)! + n \ln p + (N - n) \ln q. \quad (2)$$

The particular value $n = \bar{n}$ where P has its maximum is then determined by the condition

$$\frac{dP}{dn} = 0$$

or, equivalently, by the condition that $\ln P$ is maximum,

$$\frac{d \ln P}{dn} = \frac{1}{P} \frac{dP}{dn} = 0. \quad (3)$$

To differentiate the expression (2), we note that all the numbers occurring as factorials are large compared to unity. We can, therefore, apply to each of these the approximation (M.7) which asserts that for any number m , large enough so that $m \gg 1$,

$$\frac{d \ln m!}{dm} \approx \ln m. \quad (4)$$

Differentiation of (2) with respect to n then yields to good approximation

$$\frac{d \ln P}{dn} = -\ln n + \ln (N - n) + \ln p - \ln q. \quad (5)$$

To find the maximum of P we equate the expression (5) to zero in accordance with (3). Thus

$$\ln \left[\frac{(N - n)p}{nq} \right] = 0$$

$$\text{or} \quad \frac{(N - n)p}{nq} = 1.$$

$$\text{Hence} \quad (N - n)p = nq$$

$$\text{or} \quad Np = n(p + q).$$

Since $p + q = 1$, the value $n = \bar{n}$ where P has its maximum is then given by

$$\boxed{\bar{n} = Np.} \quad (6)$$

To investigate the behavior of $\ln P$ near its maximum, we need only expand it in a Taylor series about the value $\bar{n}$. Thus we can write

$$\ln P(n) = \ln P(\bar{n}) + \left[\frac{d \ln P}{dn} \right] y + \frac{1}{2!} \left[\frac{d^2 \ln P}{dn^2} \right] y^2 + \frac{1}{3!} \left[\frac{d^3 \ln P}{dn^3} \right] y^3 + \dots \quad (7)$$

$$\text{where} \quad y \equiv n - \bar{n} \quad (8)$$

and where the square brackets indicate that the derivatives are to be evaluated at $n = \bar{n}$. The first derivative vanishes since we are expanding about a maximum where (3) is satisfied. The other derivatives can be found by successive differentiation of (5). In particular,

$$\frac{d^2 \ln P}{dn^2} = -\frac{1}{n} - \frac{1}{N-n} = -\frac{N}{n(N-n)}.$$

Evaluating this derivative where $n = \bar{n}$, i.e., where $n = Np$ and $N - n = N(1 - p) = Nq$, we find

$$\left[\frac{d^2 \ln P}{dn^2} \right] = -\frac{1}{Npq}.$$

Hence (7) becomes

$$\ln P(n) = \ln P(\bar{n}) - \frac{y^2}{2Npq} + \dots$$

$$\text{or} \quad P(n) = \tilde{P} e^{-y^2/2Npq} \dots = \tilde{P} e^{-(n-\bar{n})^2/2Npq} \dots \quad (9)$$

where we have written $\tilde{P} \equiv P(\bar{n})$.

Note that the probability $P(n)$ in (9) becomes negligible compared to its maximum value $\tilde{P}$ when y becomes so large that $y^2/(Npq) \gg 1$ or $|y| \gg (Npq)^{1/2}$, since the exponential factor is then very much smaller than unity. The probability $P(n)$ is thus only appreciable in the domain where $|y| \lesssim (Npq)^{1/2}$. But there y is ordinarily sufficiently small so that the terms in (7) which involve y^3 and higher powers of y are negligible compared to the leading term, involving y^2 , which we retained.† Hence we can conclude that (9) is indeed a good approximation to the probability $P(n)$ in the entire region where this probability has appreciable magnitude.

† This statement is true to the extent that $(Npq)^{1/2} \gg 1$. See Prob. P. 3.

The value of the constant $\tilde{P}$ in (9) can be expressed directly in terms of p and q by using the normalization condition

$$\sum_n P(n) = 1 \quad (10)$$

where the summation is over all possible values of n . Since $P(n)$ changes only slightly between the successive integral values of n , this sum can be replaced by an integration. A range of n of magnitude dn (much larger than unity) contains dn possible values of $P(n)$. Hence the condition (10) becomes

$$\begin{aligned} \int P(n) dn &= \int_{-\infty}^{\infty} \tilde{P} e^{-(n-\bar{n})^2/2Npq} dn \\ &= P \int_{-\infty}^{\infty} e^{-y^2/2Npq} dy = 1. \end{aligned} \quad (11)$$

Here we have made the simplification of extending the range of integration from $-\infty$ to $+\infty$. This is an excellent approximation since $P(n)$ is negligibly small anyhow whenever $|n - \bar{n}|$ becomes sufficiently large. Using the relation (M.23), the integral in (11) yields simply

$$\tilde{P} \sqrt{2\pi Npq} = 1.$$

Thus

$$\tilde{P} = \frac{1}{\sqrt{2\pi Npq}}. \quad (12)$$

Using this result and the value $\bar{n} = Np$ given by (6), the expression (9) for the probability $P(n)$ then becomes

$$P(n) = \frac{1}{\sqrt{2\pi Npq}} e^{-(n-Np)^2/2Npq}. \quad (13)$$

Note that this expression is much easier to evaluate than (1) since it does not require the calculation of any factorials.

A probability of the functional form given by the right side of (9) or (13) is known as a *Gaussian distribution*. The essential argument leading to this distribution, i.e., a power series expansion of a logarithm, is widely applicable. It is therefore not surprising that Gaussian distributions occur very frequently in statistical arguments whenever the numbers under consideration are large.

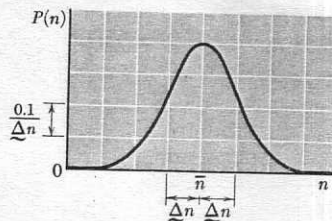


Fig. A.1 Gaussian distribution showing as a smooth curve the values of the probability $P(n)$ as a function of n . The probability $W(x)$ that n assumes a value between $\bar{n} - x$ and $\bar{n} + x$ is given by the area lying in this range below the curve. If Δn denotes the standard deviation of n , computation shows that

$$W(\Delta n) = 0.683, \quad W(2\Delta n) = 0.954,$$

and

$$W(3\Delta n) = 0.997.$$

The expression (13) for $P(n)$ can be readily used to calculate various mean values of n . The calculation of sums can be reduced to an evaluation of equivalent integrals in the same manner as that used in calculating the normalization condition (10) by means of (11). Thus one finds

$$\begin{aligned}\bar{n} &\equiv \sum_n P(n)n \\ &= (2\pi Npq)^{-1/2} \int_{-\infty}^{\infty} e^{-(n-Np)^2/2Npq} n \, dn = (2\pi Npq)^{-1/2} \int_{-\infty}^{\infty} e^{-y^2/2Npq} (\bar{n} + y) \, dy \\ &= \bar{n}(2\pi Npq)^{-1/2} \int_{-\infty}^{\infty} e^{-y^2/2Npq} \, dy + (2\pi Npq)^{-1/2} \int_{-\infty}^{\infty} e^{-y^2/2Npq} y \, dy.\end{aligned}$$

Here the first integral is the same as that occurring in (11) and has the value $(2\pi Npq)^{1/2}$. The second integral vanishes by symmetry since its integrand is odd (i.e., is of opposite sign for $+y$ and $-y$) so that contributions to the integral near $+y$ and $-y$ cancel each other. Thus we are left simply with

$$\bar{n} = \bar{n} = Np. \quad (14)$$

This shows that the mean value of n is equal to the value $\bar{n} = Np$ where the probability P is maximum.

Similarly one finds for the dispersion of n the result

$$\begin{aligned}(\Delta n)^2 &= \overline{(n - \bar{n})^2} = \sum_n P(n)(n - Np)^2 = (2\pi Npq)^{-1/2} \int_{-\infty}^{\infty} e^{-(n-Np)^2/2Npq} (n - Np)^2 \, dn \\ &= (2\pi Npq)^{-1/2} \int_{-\infty}^{\infty} e^{-y^2/2Npq} y^2 \, dy.\end{aligned}$$

Using (M.26), this integral yields

$$(\Delta n)^2 = Npq. \quad (15)$$

The standard deviation of n is then†

$$\Delta n = \sqrt{Npq}. \quad (16)$$

The Gaussian distribution (13) can thus be expressed solely in terms of the two parameters $\bar{n}$ and Δn of (14) and (16)

$$P(n) = \frac{1}{\sqrt{2\pi} \Delta n} \exp \left[-\frac{1}{2} \left(\frac{n - \bar{n}}{\Delta n} \right)^2 \right].$$

Introducing the variable

$$z \equiv \frac{n - \bar{n}}{\Delta n} \quad \text{so that} \quad n = \bar{n} + (\Delta n)z,$$

Eq. (17) can also be expressed more compactly as

$$P(n) \Delta n = \frac{1}{\sqrt{2\pi}} e^{-(1/2)z^2}.$$

Note that the Gaussian distribution is symmetric about its mean value, i.e., $P(n)$ has the same value for z and $-z$.

(17)

† Note that (14) and (15) agree properly with the results (2.66) and (2.67) derived in the text under the most general conditions of arbitrary magnitude of N .

A.2 Poisson Distribution

Consider again the binomial distribution derived in (2.14),

$$P(n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}. \quad (18)$$

In the preceding section we showed that, when $N \gg 1$, the expression (18) can be approximated by a Gaussian distribution valid in the entire domain where the probability $P(n)$ is of appreciable magnitude (i.e., where it is not too far removed from the region of its maximum). We shall now examine an approximation of (18) valid in a different domain. This approximation becomes relevant when the probability p is sufficiently small so that

$$p \ll 1 \quad (19)$$

and when the number n of interest is sufficiently small so that

$$n \ll N. \quad (20)$$

In contrast to the situation envisaged in discussing the Gaussian approximation, the number n may here be arbitrarily small.

Let us now examine the approximations made possible by the conditions (19) and (20). We note first that

$$\frac{N!}{(N-n)!} = N(N-1)(N-2) \cdots (N-n+1).$$

Since $n \ll N$, each of the n factors on the right is essentially equal to N . Hence we obtain the approximate result

$$\frac{N!}{(N-n)!} \approx N^n. \quad (21)$$

We examine next the factor

$$y \equiv (1-p)^{N-n}$$

or, equivalently, its logarithm

$$\ln y = (N-n) \ln (1-p).$$

Since $n \ll N$, we can put $N-n \approx N$. Furthermore, the condition $p \ll 1$ allows us to approximate the logarithm by the first term in a Taylor expansion, i.e., to put $\ln (1-p) \approx -p$.

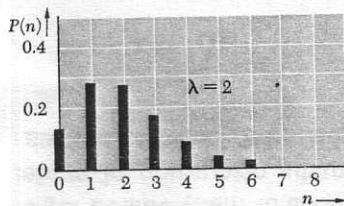
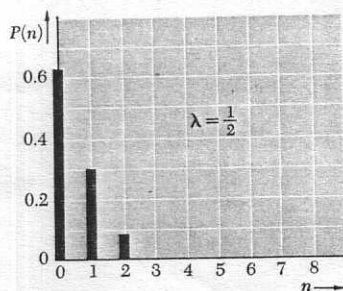


Fig. A.2 The Poisson distribution $P(n)$ of (23) as a function of n . The two cases shown correspond to a mean value $\bar{n} = \lambda$ of $\lambda = \frac{1}{2}$ and $\lambda = 2$.

Hence

$$\ln y \approx -Np$$

or

$$y \equiv (1 - p)^{N-n} \approx e^{-Np}. \quad (22)$$

Using the approximations (21) and (22) in the expression (18), we then obtain

$$P(n) = \frac{N^n}{n!} p^n e^{-Np}$$

or

$$P(n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad (23)$$

where

$$\lambda \equiv Np. \quad (24)$$

In terms of this definition of λ , the condition (19) is equivalent to

$$\lambda \ll N. \quad (25)$$

The result (23) is called the *Poisson distribution*. Note that the factor $n!$ in the denominator causes $P(n)$ to decrease very rapidly when n becomes sufficiently large. Indeed, when $\lambda < 1$, λ^n itself is a decreasing function of n so that $P(n)$ decreases monotonically as a function of n . When $\lambda > 1$, λ^n is an increasing function of n so that the factor $\lambda^n/n!$, and hence also $P(n)$, tends to exhibit a maximum near $n \approx \lambda$ before falling off for larger values of n .† In any event, when $n \gg \lambda$, the probability $P(n)$ becomes negligibly small. In the entire region where $n \leq \lambda$ so that $P(n)$ is not negligible, the condition (25) then implies that $n \leq \lambda \ll N$. The requirement (20) used in deriving the Poisson distribution is thus automatically satisfied wherever the probability $P(n)$ is of appreciable magnitude.

The parameter λ defined in (24) is, by virtue of (2.66), equal to the mean value $\bar{n}$ of n . Thus

$$\lambda = \bar{n}. \quad (26)$$

Note incidentally that, for a given value of λ or $\bar{n}$, the condition (25) or (20), requiring that $p \ll 1$, becomes increasingly well satisfied as $N \rightarrow \infty$. The Poisson distribution thus becomes always applicable in this limit.

† When N is large and $\lambda \gg 1$, the Poisson distribution (23) reduces properly to a Gaussian distribution for values of n not too far removed from λ .

Explicit verification that $\lambda = \bar{n}$

The result (26) follows also directly from the Poisson distribution (23). Using the definition of the mean value we have

$$\bar{n} \equiv \sum_{n=0}^N P(n)n = e^{-\lambda} \sum_{n=0}^N \frac{\lambda^n}{n!} n.$$

No significant error results if this sum is extended to infinity since $P(n)$ becomes negligibly small when n is large. Noting that the term with $n = 0$ vanishes, we

then obtain, putting $k = n - 1$,

$$\begin{aligned}\bar{n} &= e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{(n-1)!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{k!} \\ &= e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \lambda e^{\lambda}\end{aligned}$$

since the last sum is simply the series expansion of the exponential function.

Thus

$$\bar{n} = \lambda. \quad (27)$$

A.3 Magnitude of Energy Fluctuations

Consider two macroscopic systems A and A' in thermal interaction with each other. We shall use the notation of Sec. 4.1 and examine more closely the probability $P(E)$ that A has an energy between E and $E + \delta E$. In particular, we should like to investigate the behavior of $P(E)$ near the energy $E = \bar{E}$ where it is maximum.

For this purpose we examine the slowly varying logarithm of $P(E)$ given by (4.6),

$$\ln P(E) = \ln C + \ln \Omega(E) + \ln \Omega'(E'), \quad (28)$$

and expand it in a Taylor series about the value $\bar{E}$. Introducing the energy difference

$$\epsilon \equiv E - \bar{E}, \quad (29)$$

the Taylor series for $\ln \Omega(E)$ becomes

$$\ln \Omega(E) = \ln \Omega(\bar{E}) + \left[\frac{\partial \ln \Omega}{\partial E} \right] \epsilon + \frac{1}{2} \left[\frac{\partial^2 \ln \Omega}{\partial E^2} \right] \epsilon^2. \quad (30)$$

Here the square brackets indicate that the derivatives are all evaluated for $E = \bar{E}$. Terms involving powers of ϵ greater than ϵ^2 have been neglected. By introducing the abbreviations

$$\beta \equiv \left[\frac{\partial \ln \Omega}{\partial E} \right] \quad (31)$$

and

$$\gamma \equiv - \left[\frac{\partial^2 \ln \Omega}{\partial E^2} \right] = - \left[\frac{\partial \beta}{\partial E} \right], \quad (32)$$

Eq. (30) can be written in the simple form

$$\ln \Omega(E) = \ln \Omega(\bar{E}) + \beta\epsilon - \frac{1}{2}\gamma\epsilon^2. \quad (33)$$

The minus sign has been introduced in the definition (32) for the sake of convenience so that the parameter γ will turn out to be positive [in accordance with (4.32)].

A similar Taylor series can readily be written down for $\ln \Omega'(E')$, where $E' = E^* - E$. Expanding about the value $\bar{E}' \equiv E^* - \bar{E}$, we have

$$E' - \bar{E}' = -(E - \bar{E}) = -\epsilon.$$

Hence we obtain, analogously to (30),

$$\ln \Omega'(E') = \ln \Omega'(\bar{E}') + \beta'(-\epsilon) - \frac{1}{2}\gamma'(-\epsilon)^2 \quad (34)$$

where
$$\beta' \equiv \left[\frac{\partial \ln \Omega'}{\partial E'} \right]$$

and
$$\gamma' \equiv - \left[\frac{\partial^2 \ln \Omega'}{\partial E'^2} \right] = - \left[\frac{\partial \beta'}{\partial E'} \right]$$

are defined analogously to (31) and (32) in terms of the derivatives evaluated at $E' = \bar{E}'$. Adding (33) and (34), we then obtain

$$\ln \{\Omega(E)\Omega'(E')\} = \ln \{\Omega(\bar{E})\Omega'(\bar{E}')\} + (\beta - \beta')\epsilon - \frac{1}{2}(\gamma + \gamma')\epsilon^2. \quad (35)$$

At the value $E = \bar{E}$ where $P(E) = C\Omega(E)\Omega'(E')$ is maximum, it follows by (4.8) that $\beta = \beta'$; hence the term linear in ϵ vanishes, as it should. Thus (28) can be written in the form

$$\ln P(E) = \ln P(\bar{E}) - \frac{1}{2}\gamma_0\epsilon^2$$

or

$$P(E) = P(\bar{E})e^{-(1/2)\gamma_0(E-\bar{E})^2} \quad (36)$$

where†

$$\gamma_0 \equiv \gamma + \gamma'. \quad (37)$$

The result (36) shows that the value of γ_0 must be positive to guarantee that the probability $P(E)$ has a maximum (rather than a minimum) at $E = \bar{E}$. Indeed, it shows explicitly that $P(E)$ becomes negligibly small compared to its maximum value when $\frac{1}{2}\gamma_0(E - \bar{E})^2 \gg 1$, i.e., when $|E - \bar{E}| \gg \gamma_0^{-1/2}$. In other words, it is very improbable that

† Note that our whole argument has been similar to that used in Appendix A.1 and that (36) is indeed a Gaussian distribution.

the energy of A lies far outside the range $\bar{E} \pm \Delta E$, where†

$$\Delta E = \gamma_0^{-1/2}. \quad (38)$$

The order of magnitude of ΔE can readily be estimated by using the definition (32) of γ and the approximate expression (3.38) for $\Omega(E)$. Thus we can write for any ordinary system with a ground state energy E_0

$$\ln \Omega \sim f(E - E_0) + \text{constant}.$$

Hence we obtain by the definitions (31) and (32), evaluated for $E = \bar{E} = E$,

$$\beta = \left[\frac{\partial \ln \Omega}{\partial E} \right] \sim \frac{f}{E - E_0}$$

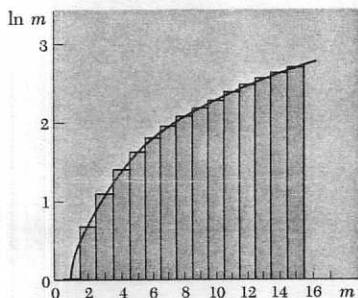
$$\text{and} \quad \gamma = - \left[\frac{\partial \beta}{\partial E} \right] \sim \frac{f}{(E - E_0)^2} \sim \frac{\beta^2}{f}. \quad (39)$$

This last relation shows explicitly that γ is positive. It also shows that, for the given value of β at which the systems are in equilibrium with each other, the smaller system (i.e., the one with the smaller number of degrees of freedom) has the larger value of γ . In (37) the magnitude of γ_0 is thus predominantly determined by the *smaller* of the two systems. Suppose, for example, that A is much smaller than A' so that $\gamma \gg \gamma'$ and $\gamma_0 \approx \gamma$. Then it follows from (38) and (39) that

$$\Delta E \sim \frac{E - E_0}{\sqrt{f}}. \quad (40)$$

Since f is a very large number in the case of a macroscopic system, Eq. (40) shows that the *relative* magnitude $\Delta E / (E - E_0)$ of energy fluctuations is very small. This result is discussed in greater detail in Sec. 4.1 where Eq. (4.10) is based on (40).

† Since (36) depends only on the absolute value $|E - \bar{E}|$ and is thus symmetric about the value $\bar{E}$, the mean value of the energy must be equal to $\bar{E}$, i.e., $E = \bar{E}$. This result is simply that already obtained in Eq. (A.17) for any Gaussian distribution. Similarly, it follows from (A.17) that ΔE in (38) is equal to the standard deviation of the energy E .

Fig. M.1 Behavior of $\ln m$ as a function of m .

M.4 Value of $\ln n!$ for large n

Since the calculation of $n!$ becomes very laborious if n is large, we should like to find a simple approximation useful for computing $n!$ in this case. By its definition

$$n! = 1 \times 2 \times 3 \times \cdots \times (n-1) \times n.$$

$$\text{Hence} \quad \ln n! = \ln 1 + \ln 2 + \cdots + \ln n = \sum_{m=1}^n \ln m. \quad (8)$$

If n is large, all terms in the sum (8) (except the first few which are also the smallest) correspond to values of m large enough so that $\ln m$ varies only slightly when m is increased by unity. The sum (8) (given by the area under the rectangles of Fig. M.1) can then be approximated with little error by an integral (giving the area under the continuous curve of Fig. M.1). With this approximation (8) becomes

$$\ln n! \approx \int_1^n \ln x \, dx = \left[x \ln x - x \right]_1^n. \quad (9)$$

Hence,

if $n \gg 1$,

$$\ln n! \approx n \ln n - n \quad (10)$$

since the contribution of the lower limit in (9) is then negligible.

A better approximation, which is good within less than one per cent error in $n!$ even if n is as small as 10, is given by *Stirling's formula*

$$\ln n! = n \ln n - n + \frac{1}{2} \ln (2\pi n). \quad (11)$$

When n is quite large, $n \gg \ln n$ so that Stirling's formula reduces to the simpler form (10).

Note also that (10) implies the result

$$\frac{d \ln n!}{dn} = \ln n + n \left(\frac{1}{n} \right) - 1 = \ln n$$

which agrees with (7).

M.5 The Inequality $\ln x \leq x - 1$

We wish to compare $\ln x$ with x itself for positive values of x . Consider the difference function

$$f(x) \equiv x - \ln x. \quad (12)$$

$$\left. \begin{array}{lll} \text{As } x \rightarrow 0, & \ln x \rightarrow -\infty; & \text{hence } f(x) \rightarrow \infty. \\ \text{As } x \rightarrow \infty, & \ln x \ll x; & \text{hence } f(x) \rightarrow \infty. \end{array} \right\} \quad (13)$$

To investigate the behavior of $f(x)$ between these limits, we note that

$$\frac{df}{dx} = 1 - \frac{1}{x} = 0 \quad \text{for } x = 1. \quad (14)$$

Since $f(x)$ is a continuous function of x satisfying (13) and having a single extremum at $x = 1$, it follows that $f(x)$ must have the appearance shown in Fig. M.2 with a minimum at $x = 1$. Hence

$$f(x) \geq f(1) \quad (= \text{sign if } x = 1),$$

or by (12),

$$x - \ln x \geq 1.$$

$$\text{Thus} \quad \ln x \leq x - 1 \quad (= \text{sign if } x = 1). \quad (15)$$

M.6 Evaluation of the Integral $\int_{-\infty}^{\infty} e^{-x^2} dx$

The *indefinite* integral $\int e^{-x^2} dx$ cannot be evaluated in terms of elementary functions. Let I denote the desired *definite* integral

$$I \equiv \int_{-\infty}^{\infty} e^{-x^2} dx. \quad (16)$$

This integral can be evaluated by exploiting the properties of the exponential function. Thus we can equally well write (16) in terms of a different variable of integration, i.e.,

$$I \equiv \int_{-\infty}^{\infty} e^{-y^2} dy. \quad (17)$$

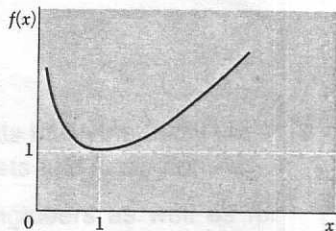


Fig. M.2 The function $f(x) \equiv x - \ln x$ as a function of x .

Multiplication of (16) and (17) then yields

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} dx dy \end{aligned}$$

or

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy. \quad (18)$$

This double integral extends thus over the entire xy plane.

Let us express the integration over this plane in terms of polar coordinates r and φ . Then one has simply $x^2 + y^2 = r^2$, and the element of area in these coordinates is given by $(r dr d\varphi)$. In order to cover the entire plane, the variables φ and r must range over the values $0 < \varphi < 2\pi$ and $0 < r < \infty$. Hence (18) becomes

$$I^2 = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r dr d\varphi = 2\pi \int_0^{\infty} e^{-r^2} r dr \quad (19)$$

since the integration over φ is immediate. But the factor r in this integrand makes the evaluation of this last integral trivial. Thus

$$I^2 = 2\pi \int_0^{\infty} \left(-\frac{1}{2}\right) d(e^{-r^2}) = -\pi \left[e^{-r^2} \right]_0^{\infty} = -\pi(0 - 1) = \pi$$

or

$$I = \sqrt{\pi}.$$

Thus

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}. \quad (20)$$

Since e^{-x^2} assumes the same value for x and $-x$, it also follows that

$$\int_{-\infty}^{\infty} e^{-x^2} dx = 2 \int_0^{\infty} e^{-x^2} dx.$$

Hence

$$\int_0^{\infty} e^{-x^2} dx = \frac{1}{2} \sqrt{\pi}. \quad (21)$$

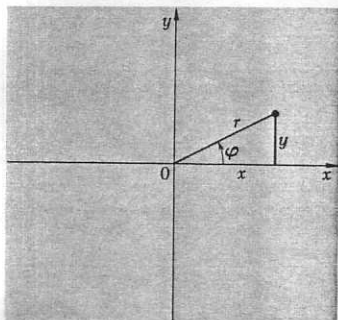


Fig. M.3 The polar coordinates r and φ used for evaluating the integral (18).

M.7 Evaluation of Integrals of the Form $\int_0^\infty e^{-\alpha x^2} x^n dx$

Denote the desired integral by

$$I_n \equiv \int_0^\infty e^{-\alpha x^2} x^n dx. \quad (22)$$

Putting $x = \alpha^{-1/2}y$, the integral becomes for $n = 0$

$$I_0 = \alpha^{-1/2} \int_0^\infty e^{-y^2} dy = \frac{\sqrt{\pi}}{2} \alpha^{-1/2} \quad (23)$$

where we have used the result (21). Similarly,

$$I_1 = \alpha^{-1} \int_0^\infty e^{-y^2} y dy = \alpha^{-1} \left[-\frac{1}{2} e^{-y^2} \right]_0^\infty = \frac{1}{2} \alpha^{-1}. \quad (24)$$

All other integrals where n is any integer such that $n \geq 2$ can then be calculated in terms of I_0 or I_1 by successive integrations by parts. Indeed,

$$\begin{aligned} \int_0^\infty e^{-\alpha x^2} x^n dx &= -\frac{1}{2\alpha} \int_0^\infty d(e^{-\alpha x^2}) x^{n-1} \\ &= -\frac{1}{2\alpha} \left[e^{-\alpha x^2} x^{n-1} \right]_0^\infty + \frac{n-1}{2\alpha} \int_0^\infty e^{-\alpha x^2} x^{n-2} dx. \end{aligned}$$

Since the integrated term vanishes at both limits, we thus obtain

$$I_n = \left(\frac{n-1}{2\alpha} \right) I_{n-2}. \quad (25)$$

For example,

$$I_2 = \frac{I_0}{2\alpha} = \frac{\sqrt{\pi}}{4} \alpha^{-3/2}. \quad (26)$$